Birkhoff Normal Form and Hamiltonian PDEs

Benoît Grébert

BIRKHOFF NORMAL FORM AND HAMILTONIAN PDES

by

Benoît Grébert

Abstract. — These notes are based on lectures held at the Lanzhou university (China) during a CIMPA summer school in july 2004 but benefit from recent devellopements. Our aim is to explain some normal form technics that allow to study the long time behaviour of the solutions of Hamiltonian perturbations of integrable systems. We are in particular interested with stability results.

Our approach is centered on the Birkhoff normal form theorem that we first proved in finite dimension. Then, after giving some exemples of Hamiltonian PDEs, we present an abstract Birkhoff normal form theorem in infinite dimension and discuss the dynamical consequences for Hamiltonian PDEs.

Résumé **(Forme normale de Birkhoff et EDP Hamiltoniènnes)**

Ces notes sont basées sur un cours donné à l'université de Lanzhou (Chine) durant le mois de juillet 2004 dans le cadre d'une école d'été organisée par le CIMPA. Cette rédaction bénéficie aussi de développements plus récents. Le but est d'expliquer certaines techniques de forme normale qui permettent d'étudier le comportement pour des temps longs des solutions de perturbations Hamiltoniènnes de systèmes intégrables. Nous sommes en particulier intéressés par des résultats de stabilité.

Notre approche est centrée sur le théorème de forme normale de Birkhoff que nous rappelons et démontrons d'abord en dimension finie. Ensuite, après avoir donné quelques exemples d'EDP Hamiltoniènnes, nous démontrons un théorème de forme normale de Birkhoff en dimension infinie et nous en discutons les applications à la dynamique des EDP Hamiltoniènnes.

Key words and phrases. — Birkhoff normal form, Resonances, KAM theory, Hamiltonian PDEs, long time stability.

Contents

1. Introduction

The class of Hamiltonian systems close to integrable system contain most of the important physic models. Typically a Hamiltonian system in finite dimension reads (cf. section 2)

$$\begin{cases} \dot{q}_j = \frac{\partial H}{\partial p_j}, \quad j = 1, \dots, n \\ \dot{p}_j = -\frac{\partial H}{\partial q_j}, \quad j = 1, \dots, n \end{cases}$$

where the Hamiltonian H is a smooth fonction from $\mathbb{R}^{2N}$ to $\mathbb{R}$. In these lectures we are interesred in the case where H decomposes in $H = H_0 + \epsilon P$, H_0 being integrable in the sense that we can "integrate" the Hamiltonian system associated to H_0 (cf. section 2.3), P being the perturbation and ϵ a small parameter. This framework contains a lot of important example of the classical mechanics. If we allow the number of degree of freedom, N, to grow to infinity, then we arrive in the world of quantum mechanics and the corresponding equations are typically nonlinear partial differential equations (PDEs). Again a lot of classical examples are included in this framework like, for instance, the nonlinear wave equation, the nonlinear Schrödinger equation or the Korteweg-de Vries equation (cf. section 5.1).

The historical example (in finite dimension) is given by the celestial mechanics: More than 300 years ago Newton gaves the evolution equation for a system of N heavy bodies under the action of the gravity.
When $N = 2$, Kepler gaves the solution, the bodies describe ellipses. Actually for $N = 2$ the system is integrable.
As soon as $N \geq 3$ the system leaves the integrable world and we do not know the expression of the general solution. Nevertheless if we consider the celestial system composed by the Sun (S), the Earth (E) and Jupiter

(J) and if we neglect the interaction between J and E, then the system is again integrable and we find quasiperiodic solution. Mathematically the solutions read $t \mapsto g(\omega_1 t, \omega_2 t, \omega_3 t)$ where g is a regular function from the torus $T^3 = S^1 \times S^1 \times S^1$ to $\mathbb{R}^{18}$ (three postions and three moments in $\mathbb{R}^3$) and ω_j, $j = 1, 2, 3$ are frequencies. Visually J and E turn around S which turns around the center of mass. Notice that the trajectory (or orbit) is contained in the torus $g(\mathcal{T}^3)$ of dimension 3 and that this torus is invariant under the flow. On the other hand, if $(\omega_1, \omega_2, \omega_3)$ are rationnaly independent, then the trajectory densely fills this torus while, if for instance the three frequencies are rationnally proportional, then the trajectoty is periodic and describes a circle included in $g(\mathcal{T}^3)$.
Now the exact system S-E-J is described by a Hamiltonian $H = H_0 + \epsilon P$ in which H_0 is the integrable Hamiltonian where we neglect the interaction E-J, P takes into account this interaction and $\epsilon = \dfrac{\text{jupiter's mass + earth's mass}}{\text{sun's mass}}$ plays the rule of the small parameter.
Some natural questions arrive:

- Do invariant tori persist after this small perturbation?
- At least are we able to insure stability in the sense that the planets remain in a bouded domain?
- Even if we are unable to answer these questions for eternity, can we do it for very large -but finite- times?

These questions have interested a lot of famous mathematicians and physicists. In the 19-th century one tried to expand the solutions in perturbative series: $u(t) = u_0(t) + \epsilon u_1(t) + \epsilon^2 u_2(t) + \ldots$, the term u_{k+1} being determined by an equation involving $u_0, \ldots, u_k$. Unfortunatly this series does not converge. This convergence problems seemed so involved that, at the principle of the 20-th century, most of scientist believed in the *ergodic hypothesis*: typically, after arbitrarily small perturbation, all the trajectories fill all the phase space and the stable trajectories are exceptionnal. Actually, H. Poinaré proved that a dense set of invariant tori are destroyed by an arbitrarily small perturbation. Nevertheless, a set can be dense but very small and in 1954 A. N. Kolmogorov [**Kol54**] announced that the majority (in the measure sense) of tori survive (see section 7). The proof of this result was completed by V. Arnold [**Arn63**] and J. Moser [**Mos62**] giving birth to the KAM theory.

In order to illustrate this result we can apply it to a simplified S-E-J system: we assume that the S-E-J system reduces to a Hamiltonian system with

3 degrees of freedom without symmetries (the symmetries of the true system complicates the pictures and generates degenerancies). In this case, the KAM theorem says, roughly speaking (see theorem 7.4 for a precise statement), that if $(\omega_1, \omega_2, \omega_3) \in \mathcal{C}$, a Cantor set of $\mathbb{R}^3$ having a positive measure, or equivalently if the initial positions and moments are in a Cantor set, then the trajectory is quasi periodic. Since a Cantor set has an empty interior the condition $(\omega_1, \omega_2, \omega_3) \in \mathcal{C}$ is not physical (no measurement could decide if this condition is verified or not).
The present lectures will be centered on the Birkhoff normal form approach which does not control the solution for any times but does not require an undecidable hypothesis. In the case of our simplified S-E-J system, the Birkhoff normal form theorem says, roughly speaking, that having fixed an integer $M \geq 1$, and $\epsilon < \epsilon_0(M)$ small enough, to any initial datum corresponding to not rationaly dependent frequencies $(\omega_1, \omega_2, \omega_3)$, we can associate a torus such that the solution remains ϵ-close to that torus during a lapse of time greater than $1/\epsilon^M$ (see section 3 for a precise statement). Note that this result can be physically sufficient if $1/\epsilon^M$ is greater than the age of the universe.

The rational independence of the frequencies (one also says the nonresonancy) is of course essential in all this kind of perturbative theorems. Again we can illustrate this fact with our system S-E-J: suppose that, when considering the system without E-J interaction, the three bodies are periodically align, the Earth being between Jupiter and the Sun (notice that this implies that the frequencies $(\omega_1, \omega_2, \omega_3)$ are rationaly dependent). When we turn on the interaction E-J, Jupiter will attract the Earth outside of its orbit periodically (i.e. when the three bodies are align or almost align), these accumulate small effects will force the earth to escape its orbit and thus the invariant torus will be destroyed.

The generalisation of these results to the infinite diemensional case is of course not easy but it worth trying: The expected results may apply to nonlinear PDEs when they can be viewed as an infinite dimensional Hamiltonian system (cf. section 5) and concern the long time behaviour of the solution, a very difficult and competitive domain.

For a general overview on Hamiltonian PDEs, the reader may consult the recent monographies by Craig [**Cra00**], by Kuksin [**Kuk00**], by Bourgain [**Bou05a**] and by Kappleler and Pöschel [**KP03**]. In the present lectures we mainly focus on the extension of the Birkhoff normal form theorem. Such extension was first (partially) achieved by Bourgain [**Bou96**] and then by

Bambusi [**Bam03**]. The results stated in this text was first proved by Bambusi and myself in [**BG04**]. The proof presented here and some generalisations benefit of a recent collaboration with Delort and Szeftel [**BDGS05**].

After this general presentation, I give a brief outline of the next sections:

– **Section 2** : Hamiltonian formalism in finite dimension.

  We recall briefly the classical Hamiltonian formalism including: integrals of the motion, Lie transforations, Integrability in the Liouville sense, action angle variables, Arnold-Liouville theorem (see for instance [**Arn89**] for a complete presentation).

– **Section 3** : The Birkhoff normal form theorem in finite dimension.

  We state and prove the Birkhoff normal form theorem and then present its dynamical consequences. Theses results are well known and the reader may consult [**MS71, HZ94, KP03**] for more details and generalizations.

– **Section 4** : A Birkhoff normal form theorem in infinite dimension.

  We state a Birkhoff normal form theorem in infinite dimension and explain its dynamical consequences. In particular, results on the longtime behaviour of the solutions are discussed. This is the most important part of this course. A slightly more general abstract Birkhoff theorem in infinite dimension was obtained in [**BG04**] and the dynamical consequences was also obtained there.

– **Section 5** : Application to Hamiltonian PDEs.

  Two examples of Hamiltonian PDEs are given: the nonlinear wave equation and the nonlinear Schrödinger equation. We then verify that our Birkhoff theorem and its dynamical consequences apply to both examples.

– **Section 6** : Proof of the Birkhoff normal form theorem in infinite dimension.

  Instead of giving the proof of [**BG04**], we present a simpler proof using a class of polynomials first introduced in [**DS04**], [**DS05**]. Actually we freely used notations and parts of proofs of theses three references.

– **Section 7** : Generalisations and comparison with KAM type results.

  In a first part we comment on some generalisations of our result. In the second subsection, we try to give to the reader an idea on the KAM theory in both finite and infinite dimension. Then we compare the Birkhoff approach with the KAM approach.

Acknowledgements: it is a great pleasure to thank D. Bambusi and J. M. Delort for helpful discussions on these notes.

2. Hamiltonian formalism in finite dimension

2.1. Basic definitions. — We only consider the case where the **phase space** (or configuration space) is an open set, M, of $\mathbb{R}^{2n}$. We denote by J the canonical **Poisson matrix**, i.e.

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} .$$

More generally, J could be an antisymmetric matrix on $\mathbb{R}^{2n}$. All the theory can be extended to the case where the phase space is a $2n$ dimensional symplectic manifold.
A **Hamiltonian fonction**, H, is a regular real valued function on the phase space, i.e. $H \in C^\infty(M, \mathbb{R})$. To H we associate the **Hamiltonian vector field**

$$X_H(q,p) = J\nabla_{q,p}H(q,p)$$

where $\nabla_{p,q}H$ denotes the gradient of H with respect to p, q, i.e.

$$\nabla_{q,p}H = \begin{pmatrix} \frac{\partial H}{\partial q_1} \\ \vdots \\ \frac{\partial H}{\partial q_n} \\ \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial p_n} \end{pmatrix} , \quad X_H = \begin{pmatrix} \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial p_n} \\ -\frac{\partial H}{\partial q_1} \\ \vdots \\ -\frac{\partial H}{\partial q_n} \end{pmatrix} .$$

The associated **Hamiltonian system** then reads

$$\frac{d}{dt}\begin{pmatrix} q \\ p \end{pmatrix} = X_H(q,p)$$

or equivalently

$$\begin{cases} \dot q_j = \frac{\partial H}{\partial p_j}, & j = 1, \dots, n\,, \\ \dot p_j = -\frac{\partial H}{\partial q_j}, & j = 1, \dots, n\,. \end{cases}$$

The **Poisson bracket** of two Hamiltonian functions F, G is a new Hamiltonian function $\{F, G\}$ given by

$$\{F,G\}(q,p) = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(q,p)\frac{\partial G}{\partial p_j}(q,p) - \frac{\partial F}{\partial p_j}(q,p)\frac{\partial G}{\partial q_j}(q,p) .$$

2.2. A fundamental example: the harmonic oscillator. — Let $M = \mathbb{R}^{2n}$ and

$$H(q,p) = \sum_{j=1}^{n} \omega_j \frac{p_j^2 + q_j^2}{2}$$

where

$$\omega = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix} \in \mathbb{R}^n$$

is the frequencies vector. The associated system is the **harmonic oscillator** whose equations read

$$\begin{cases} \dot{q}_j = \omega_j p_j, & j = 1, \dots, n \\ \dot{p}_j = -\omega_j q_j, & j = 1, \dots, n \end{cases}$$

and whose solutions are quasi-periodic functions given by

$$\begin{cases} q_j(t) = q_j(0)\cos\omega_j t + p_j(0)\sin\omega_j t, & j = 1, \dots, n \\ p_j(t) = -q_j(0)\sin\omega_j t + p_j(0)\cos\omega_j t, & j = 1, \dots, n. \end{cases}$$

Let us notice that for each j, (q_j, p_j) describes a circle of radius $\frac{p_j(0)^2 + q_j(0)^2}{2} =: I_j$ and thus the orbits of the harmonic oscillator are included in tori

$$T_I := \{(q,p) \in \mathbb{R}^{2n} \mid (p_j^2 + q_j^2)/2 = I_j, \ j = 1, \dots, n\}$$

whose dimension is generically n (it can be less if $p_j(0)^2 + q_j(0)^2 = 0$ for some j). To decide wether the orbit fills the torus or not we need the following definition:

Definition 2.1. — *A frequencies vector, $\omega \in \mathbb{R}^n$, is* ***non resonant*** *if*

$$k \cdot \omega := \sum_{j=1}^{n} k_j \omega_j \neq 0 \quad \textit{for all} \quad k \in \mathbb{Z}^n \setminus \{0\}.$$

From number theory we learn that if ω is non resonant (or not rationnally dependent) then $\{k \cdot \omega \mid k \in \mathbb{Z}^n\}$ is dense in $\mathbb{R}^n$ and thus we deduce that the orbit (or trajectory) is dense in the torus. On the contrary, if ω is resonant then the orbit is not dense in T_I but in a torus of smaller dimension. For instance if all the frequencies are rationally proportional, $k_1\omega_1 = k_2\omega_2 = \ldots = k_n\omega_n$ for some choice of $k_1, \dots, k_n$ in $\mathbb{Z}$, the orbit is a circle and the solution is in fact periodic.

2.3. Integrability. —

Definition 2.2. — *A **constant of motion** (or an integral of motion) for H is a regular function, $F \in C^\infty(M, \mathbb{R})$ satisfying $\{F, H\} = 0$.*

Proposition 2.3. — *Let $F \in C^\infty(M, \mathbb{R})$ then, if $t \mapsto (q(t), p(t))$ is a solution of the Hamiltonian system associated to H,*

$$\frac{d}{dt}F(q(t), p(t)) = \{F, H\}(q(t), p(t)).$$

In particular, if F is a constant of motion, then $F(q, p)$ is invariant under the flow generated by H.

Proof. — By definition,

$$\frac{d}{dt}F(q(t), p(t)) = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}\dot{q}_j + \frac{\partial F}{\partial p_j}\dot{p}_j = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial F}{\partial p_j}\frac{\partial H}{\partial q_j} = \{F, H\}.$$

□

In the case of the harmonic oscillator the actions I_j, $j = 1, \dots, n$, defined by

$$I_j = \frac{p_j^2 + q_j^2}{2}$$

are integrals of the motion : $\dot{I}_j = 0$.

Definition 2.4. — *A $2n$-dimensional Hamiltonian system is **integrable in the sense of Liouville** if there exist n regular functions $F_1, F_2, \dots, F_n \in C^\infty(M, \mathbb{R})$ such that*

(i) $\{F_j, H\} = 0$ *for $j = 1, \dots, n$ (i.e. the F_j are integrals of the motion).*
(ii) $\{F_j, F_k\} = 0$ *for $j, k = 1, \dots, n$ (i.e. the F_j are in involution).*
(iii) $(\nabla_{q,p}F_j)_{j=1,\dots,n}$ *are lineary independent.*

In view of this definition, the harmonic oscillator is integrable in the sense of Liouville on the dense open subset where none of the actions I_j vanishes : it suffices to choose $F_j = I_j$.
However it is not Liouville integrable on the whole space: the actions are not independent everywhere. More generally, a $2n$-dimensional Hamiltonian system which admits n integrals in involution that are independent on a dense open subset is often called a **Birkhoff integrable Hamiltonian systems**.

Examples of Liouville integrable system are obtained when the Hamiltonian depends only on p: $H(q, p) = h(p)$. In this case, often called the canonical example of integrable Hamiltonian system, $\dot{p}_j = 0$ and $(p_j)_{j=1,\dots,n}$ satisfies (i),

(ii) and (iii) of the previous definition. Actually the motion is trivial since $\dot{q}_j = \frac{\partial H}{\partial p_j}$ =constant=: ω_j and thus we can *integrate* the equations to obtain

$$\begin{cases} q_j(t) = q_j(0) + \omega_j t, & j = 1, \dots, n \\ p_j(t) = p_j(0), & j = 1, \dots, n. \end{cases}$$

Let H be a Liouville integrable Hamiltonian and denote by $F_1, F_2, \dots, F_n$ a complete set of independent integrals in involution on the phase space M. The leaves

$$M_c = \{(q,p) \in M \mid F_j(q,p) = c_j,\ j = 1, \dots, n\}$$

are smooth submanifolds of M of dimension and codimension n (1), and the whole manifold M is foliated into these leaves.

Theorem 2.5. — *(Arnold-Liouville Theorem) Let H be a Liouville integrable Hamiltonian on M a symplectic manifold of dimension $2n$. If one of its leaves is compact and connected then there exists a neighborhood $\mathcal{U}$ of this leave, a neighborhood D of 0 in $\mathbb{R}^n$ and a change of variable*(2) $\Psi : D \times T^n \ni (I, \theta) \mapsto (q,p) \in \mathcal{U}$ *such that*

(i) *$H \circ \Psi = h(I)$ is a function of the actions alone*
(ii) *the Hamiltonian formalism is preserved, i.e., in the new variables, the equations read $\dot{I}_j = 0$, $\dot{\theta}_j = -\frac{\partial h}{\partial I_j}$, $j = 1, \dots, n$ (i.e. the change of variable is a canonical transformation in the sense of the definition 2.6 below).*

That means that, every Liouville integrable Hamiltonian system with compact leaves is equivalent to the canonical one. The new variables are called the **action-angle variables**.

In the case of the harmonic oscillator, the action-angle variables are given by the symplectic polar coordinates :

$$I_j = \frac{p_j^2 + q_j^2}{2}, \quad \theta_j = \arctan \frac{q_j}{p_j},$$

they are well defined on the dense open subset where none of the actions I_j vanishes.

Notice that the Arnold-Liouville theorem implies that all the leaves M_c are tori. Therefore, in this case, the whole phase space M is foliated by invariant tori of dimension n (so called Lagrangian tori). This is not true in the case of

(1) Actually they are Lagrangian submanifolds : submanifolds of maximal dimension such that the restriction of the symplectic form to it vanishes.

(2) here $T^n = S^1 \times \ldots \times S^1$, n times, is the n dimensional torus

a Birkhoff integrable Hamiltonian system where the dimension of the leaves may vary (as in the case of the harmonic oscillator).

2.4. Perturbation of integrable Hamiltonian system. — We consider a Hamiltonian function $H = H_0 + P$ where H_0 is integrable and P is a perturbation term.

The general philosophy will consist in transforming H in such a way that the new Hamiltonian be closer to an integrable one: $H \to \tilde{H} = \tilde{H}_0 + \tilde{P}$ with $\tilde{H}_0$ still integrable and $\tilde{P} \ll P$. The first question is: How to transform H? We cannot use all changes of variable because we want to conserve the Hamiltonian structure.

Definition 2.6. — *A map $\varphi : M \ni (q,p) \mapsto (\xi,\eta) \in M$ is a* ***canonical transformation*** *(or a* ***symplectic change of coordinates****) if*

- *φ is a diffeomorphism*
- *φ preserves the Poisson Bracket : $\{F,G\}\circ\varphi = \{F\circ\varphi, G\circ\varphi\} for any * F$ and G.*

As a consequence, if $\tilde{H} = H \circ \varphi^{-1}$ with φ canonical, then the Hamiltonian system reads in the new variables (ξ,η) as in the old ones

$$\dot{\xi}_j = \frac{\partial \tilde{H}}{\partial \eta_j}\ , \quad \dot{\eta}_j = -\frac{\partial \tilde{H}}{\partial \xi_j} \quad j = 1, \dots, n.$$

There exists a very convenient way of constructing canonical transformation: Let $\chi : M \to \mathbb{R}$ a regular function and denote φ_t the flow generated by X_χ. If φ_t is well defined up to $t = 1$, the map $\varphi := \varphi_1$ is called the Lie transform associated to χ. More explicitely, the new couple of variables $(\xi,\eta) = \varphi(q,p)$ is the value at time 1 of the solution of the system $\frac{d}{dt}\begin{pmatrix}\xi\\ \eta\end{pmatrix} = X_\chi(\xi,\eta)$ whose value at $t = 0$ is (q,p). Notice that, since the map $(t;q,p) \mapsto \varphi_t(q,p)$ is defined on an open set (cf. the Cauchy-Lipschitz theorem), if φ_1 is defined at the point (q,p) then it is locally defined around (q,p).

Proposition 2.7. — *A Lie transform is canonical.*

Proof. — : classical (see for instance [**Arn89**]).

The following lemma will be essential to use the Lie transforms

Lemma 2.8. — *Let $G : M \to \mathbb{R}$ be a regular function. Then*

$$\frac{d}{dt}(G \circ \varphi_t) = \{G, \chi\} \circ \varphi_t.$$

Proof. —

$$\begin{aligned}\frac{d}{dt}(G\circ\varphi_t)(q,p) &= \nabla G(\varphi_t(q,p))\cdot\dot\varphi_t(q,p)\\ &= \nabla G(\varphi_t(q,p))\cdot J\nabla\chi(\varphi_t(q,p))\\ &= \{G,\chi\}(\varphi_t(q,p)).\end{aligned}$$

□

Then, using the Taylor expansion of $G\circ\varphi_t$ at $t=0$, evaluated at $t=1$, we obtain for any $k\geq 0$,

$$G\circ\phi(q,p)=\sum_{l=0}^{k}G_l(q,p)+(k+1)\int_0^1(1-t)^kG_{k+1}\circ\varphi_t(q,p)dt$$

with $G_l=1/l\{G_{l-1},\chi\}$ for $l\geq 1$ and $G_0=G$.

3. The Birkhoff normal form theorem in finite dimension

In this section we consider perturbations of Hamiltonian systems near an elliptic fixed point.

Let H be a Hamiltonian function on M having an isolated equilibrium. Without lost of generality we can assume that the equilibrium is at the origin in $\mathbb{R}^{2n}$, that the origin belongs to M and that $H(0,0)=0$. Then the Hamiltonian reads

$$H=\frac{1}{2}<A(q,p),(q,p)>+\text{ cubic terms}+\dots$$

where A is the Hessian of H at 0, a symmetric $2n\times 2n$ real matrix. Since we suppose the equilibrium is elliptic, the spectrum of the linearized system $\dot u=JAu$ is purely imaginary:

$$\text{spec}(JA)=\{\pm i\omega_1,\dots,\pm i\omega_n\}$$

with $\omega_1,\dots\omega_n$ real. It turns out that there exists a linear symplectic change of coordinates that brings the quadratic part of H into the following normal form (cf. [**HZ94**], section1.7, theorem 8)

$$<A(q,p),(q,p)>=\sum_{j=1}^{n}\omega_j(p_j^2+q_j^2)$$

where, for simplicity, we denote the new coordinates by the same symbols.

Therfore, in this section, we will focus on the perturbation of the harmonic oscillator

$$H_0(q,p) = \sum_{j=1}^{n} \omega_j \frac{p_j^2 + q_j^2}{2} = \sum_{j=1}^{n} \omega_j I_j, \tag{3.1}$$

where we denote $I_j(q,p) := \frac{p_j^2+q_j^2}{2}$ the j^{th} action of (q,p).
The total Hamiltonian reads $H = H_0 + P$ where P is a regular real valued function, $P \in C^\infty(M,\mathbb{R})$, which is at least cubic, $P = O(\|(q,p)\|^3)$, in such a way that, in a small neighborhood of $(0,0)$, P will appear as a perturbation of H_0.

We say that P is in **normal form** with respect to H_0 if it commutes with the integrable part: $\{P, H_0\} = 0$.
For $k \in \mathbb{Z}^n$, we denote by $|k|$ the length of k: $|k| = |k_1| + \ldots + |k_n|$. We need a refined version of the nonresonancy definition (compare with definition 2.1):

Definition 3.1. — *Let $r \in \mathbb{N}$. A frequencies vector, $\omega \in \mathbb{R}^n$, is **non resonant up to order** r if*

$$k \cdot \omega := \sum_{j=1}^{n} k_j \omega_j \neq 0 \quad \text{for all} \quad k \in \mathbb{Z}^n \text{ with } 0 < |k| \leq r.$$

Of course if ω is non resonant then it is nonresonant up to any order $r \in \mathbb{N}$.

3.1. The theorem and its dynamical consequences. — We begin stating the classical Birkhoff normal form theorem (see for instance [**Mos68, HZ94**]).

Theorem 3.2. — ***[Birkhoff Normal Form Theorem]*** *Let $H = H_0 + P$, H_0 being the harmonic oscillator (3.1) and P being a C^∞ real valued function having a zero of order 3 at the origin and fix $r \geq 3$ an integer. There exists $\tau : \mathcal{U} \ni (q',p') \mapsto (q,p) \in \mathcal{V}$ a real analytic canonical transformation from a neighborhood of the origin to a neighborhood of the origin which puts H in normal form up to order r i.e.*

$$H \circ \tau = H_0 + Z + R$$

with

(i) *Z is a polynomial or order r and is in **normal form**, i.e.: $\{Z, H_0\} = 0$.*
(ii) *$R \in C^\infty(M,\mathbb{R})$ and $R(q',p') = O(\|(q',p')\|)^{r+1}$.*
(iii) *τ is close to the identity: $\tau(q',p') = (q',p') + O(\|(q',p')\|)^2$.*

In particular if ω is non resonant up to order r then Z depends only on the new actions: $Z = Z(I'_1, \dots, I'_n)$ with $I'_j = \frac{(p'_j)^2+(q'_j)^2}{2}$.

Before proving this theorem, we analyse its dynamical consequences in the non resonant case.

Corollary 3.3. — *Assume ω is non resonant. For each $r \geq 3$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that if $\|(q_0, p_0)\| = \varepsilon < \varepsilon_0$ the solution $(q(t), p(t))$ of the Hamiltonian system associated to H which takes value (q_0, p_0) at $t = 0$ satisfies*

$$\|(q_t, p_t)\| \leq 2\varepsilon \quad \textit{for } |t| \leq \frac{C}{\varepsilon^{r-1}}.$$

Furthermore for each $j = 1, \dots, n$

$$|I_j(t) - I_j(0)| \leq \varepsilon^3 \quad \textit{for } |t| \leq \frac{C}{\varepsilon^{r-2}},$$

where $I_j(t) = I_j(q_t, p_t)$.

Proof. — Denote $z := (q, p)$ and $z' = \tau^{-1}(z)$ where τ is the transformation given by theorem 3.2. Then we define $N(z) := \|z\|^2 = 2\sum_{j=1}^n I_j(q, p)$. Using that Z depends only on the new actions, we have

$$\{N, H\}(z) = \{N \circ \tau, H \circ \tau\} \circ \tau^{-1}(z) = \{N \circ \tau, R\}(z') = O(\|z'\|^{r+1}) = O(\|z\|^{r+1})$$

Therefore $|\dot{N}| \leq CN^{(r+1)/2}$. Using that the solution to the ODE $\dot{x} = ax^d$ is given by $(d > 1)$

$$x(t) = x_0(1 - x_0^{d-1}(d-1)at)^{\frac{-1}{d-1}}$$

one easily deduces the first part of the corollary.

To prove the second part, write (with $I' = I \circ \tau^{-1}$)

$$|I_j(t) - I_j(0)| \leq |I_j(t) - I'_j(t)| + |I'_j(t) - I'_j(0)| + |I'_j(0) - I_j(0)|.$$

The first and the third term of the right side of this inequality are estimated by $c\varepsilon^3$ because $\|z - z'\| \leq c\|z\|^2$ and $|I' - I| \leq \|z' - z\|\,\|z + z'\|$. To estimate the middle term we notice that

$$\frac{d}{dt}I'_j = \{I'_j, H\} = \{I_j, H \circ \tau\} \circ \tau^{-1} = O(\|z'\|^{r+1}) \tag{3.2}$$

and therefore, for $|t| \leq \frac{c}{\varepsilon^{r-2}}$,

$$|I'_j(t) - I'_j(0)| \leq c'\varepsilon^3.$$

□

We can also prove that the solution remains close to a torus for a long time, that is the contain of the following

Corollary 3.4. — *Assume ω is non resonant. For each $r \geq 3$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that if $\|(q_0, p_0)\| = \varepsilon < \varepsilon_0$ then there exists a torus $\mathcal{T}_0 \subset M$ satisfying*

$$dist((q(t), p(t)), \mathcal{T}_0) \leq C\varepsilon^{r_1} \quad for\ |t| \leq 1/\varepsilon^{r_2}$$

where $r_1 + r_2 = r + 1$.

Proof. — Let

$$\mathcal{T}_0 := \{(q, p) \mid I_j(\tau^{-1}(q, p)) = I_j(\tau^{-1}(q_0, p_0)),\ j = 1, \dots, n\}.$$

Using (3.2) we deduce that for $|t| \leq 1/\varepsilon^{r_2}$

$$|I'_j(t) - I'_j(0)| \leq c'\varepsilon^{r_1},$$

where as before $I' = I \circ \tau^{-1}$. Therefore, using assertion (iii) of theorem 3.2 we obtain the thesis. □

Remark 3.5. — **An extension to the Nekhoroshev's theorem**
If ω is non resonant at any order, it is natural to try to optimize the choice of r in theorem 3.2 or its corollaries. Actually, if we assume that ω satisfies a diophantine condition

$$|k \cdot \omega| \geq \gamma |k|^{-\alpha} \text{ for all } k \in \mathbb{Z}^n \setminus \{0\},$$

standart estimates (see for instance [**BGG85, GG85, Pös93**]) allow to prove that for (q, p) in B_ρ, the ball centered at the origin and of radius ρ, the remainder term in theorem 3.2 is of order $(r!)^{\tau+1}\rho^{r+1}$. This leads to show that the constant C in corollary 3.3 is of order $(r!)^{-(\tau+1)}$. Namely one proves that if $\|(q_0, p_0)\| = \varepsilon$ is small enough

$$\|(q_t, p_t)\| \leq 2\varepsilon \quad \text{for } |t| \leq \frac{C}{\varepsilon^{r-1}(r!)^{\tau+1}}$$

where the new constant C depends only on P and on the number of degrees of freedom[(3)] n. Using the Stirling's formula for $r!$ and choosing $r = e^2\varepsilon^{-1/(\tau+1)}$, one obtains that the solution remains controled by 2ε during an **exponentially long time**:

$$\|(q_t, p_t)\| \leq 2\varepsilon \quad \text{for } |t| \leq C \exp\left(\frac{\beta}{\varepsilon^{1/(\tau+1)}}\right),$$

where β is a non negative constant. This last statement is a formulation of the Nekhoroshev's theorem (see [**Nek77**]).

[(3)] This dependence with respect to n makes impossible to generalize, at least easily, this remark in the infinite dimensional case.

3.2. Proof of the Birkhoff normal form theorem. — We prefer to use the complex variables

$$\xi_j = \frac{1}{\sqrt{2}}(q_j + ip_j), \quad \eta_j = \frac{1}{\sqrt{2}}(q_j - ip_j)$$

because the calculus are easier in this framework. Notice in particular that the actions read $I_j = \xi_j \eta_j$ and thus it is very simple to express that a monomial $\xi_{j_1} \dots \xi_{j_k} \eta_{l_1} \dots \eta_{l_{k'}}$ depends only on the actions, it suffices that $k = k'$ and $\{j_1, \dots, j_k\} = \{l_1, \dots, l_k\}$.

We have $H_0 = \sum_{j=1}^{n} \omega_j \xi_j \eta_j$ and we easily verify that, in these variables, the Poisson bracket reads

$$\{F, G\} = i \sum_{j=1}^{n} \frac{\partial F}{\partial \xi_j} \frac{\partial G}{\partial \eta_j} - \frac{\partial F}{\partial \eta_j} \frac{\partial G}{\partial \xi_j} \ .$$

We will say that a function F defined in the variable (ξ, η) is **real** when $F(\xi, \bar{\xi})$ is real which means that in the original variables (q, p), F is real valued.

We now begin the proof of theorem 3.2. Having fixed some $r \geq 3$, the idea is to construct iteratively for $k = 2, \dots, r$, a canonical transformation τ_k, defined on a neighborhood of the origin in M, and real functions Z_k, P_{k+1}, R_{k+2} such that

$$H_k := H \circ \tau_k = H_0 + Z_k + P_{k+1} + R_{k+2} \tag{3.3}$$

and with the following properties

(i) Z_k is a polynomial of degree k having a zero of degree 3 at the origin and $\{Z_k, H_0\} = 0$.
(ii) P_{k+1} is a homogeneous polynomial of degree $k + 1$.
(iii) R_{k+2} is a regular Hamiltonian having a zero of order $k + 2$ at the origin.

Then (3.3) at order r proves theorem 3.2 with $Z = Z_r$ and $R = P_{r+1} + R_{r+2}$.

First remark that the Hamiltonian $H = H_0 + P$ has the form (3.3) with $k = 2$, $\tau_2 = I$, $Z_2 = 0$, P_3 being the Taylor's polynomial of P at degree 3 and $R_4 = P - P_3$. We show now how to pass from k to $k + 1$.

We search for τ_{k+1} of the form $\tau_k \circ \phi_{k+1}$, ϕ_{k+1} being a Lie transform associated to the Hamiltonian function χ_{k+1}. Recall from section 2.4 that for regular F

$$F \circ \phi_{k+1} = F + \{F, \chi_{k+1}\} + 1/2\{\{F, \chi_{k+1}\}, \chi_{k+1}\} + \dots$$

We search for χ_{k+1} as a homogeneous real polynomial of degree $k+1$ and we decompose $H_k \circ \phi_{k+1}$ as follows

$$H_k \circ \phi_{k+1} = H_0 + Z_k + \{H_0, \chi_{k+1}\} + P_{k+1} \tag{3.4}$$

$$+ R_{k+2} \circ \phi_{k+1} + H_0 \circ \phi_{k+1} - H_0 - \{H_0, \chi_{k+1}\} \tag{3.5}$$

$$+ Z_k \circ \phi_{k+1} - Z_k + P_{k+1} \circ \phi_{k+1} - P_{k+1} . \tag{3.6}$$

Notice that if F_1 is a homogeneous polynomial of degree d_1 and F_2 is a homogeneous polynomial of degree d_2 then $\{F_1, F_2\}$ is a homogeneous polynomial of degree $d_1 + d_2 - 2$. Notice also that, since $\chi_{k+1}(\xi, \eta) = O(\|(\xi, \eta)\|)^{k+1}$, we have

$$\phi_{k+1}(\xi, \eta) = (\xi, \eta) + O(\|(\xi, \eta)\|)^k.$$

Using these two facts we deduce that (3.5) and (3.6) are regular Hamiltonians having a zero of order $k+2$ at the origin. Therefore, using the Taylor formula, the sum of these terms decomposes in $P_{k+2} + R_{k+3}$ with P_{k+2} and R_{k+3} satisfying the properties (ii) and (iii). So it remains to prove that χ_{k+1} can be choosen in such a way that $Z_{k+1} := Z_k + \{H_0, \chi_{k+1}\} + P_{k+1}$ satisfies (i). This is a consequence of the following lemma

Lemma 3.6. — *Let Q be a homogeneous real polynomial of degree k, there exist two homogeneous real valued polynomials χ and Z of degree k such that*

$$\{H_0, \chi\} + Q = Z \tag{3.7}$$

and

$$\{Z, H_0\} = 0. \tag{3.8}$$

Equation (3.7) is known in the literature as the **homological equation**.

Proof. — For $j \in [1, \dots, n]^{k_1}$ and $l \in [1, \dots, n]^{k_2}$, denote $\xi^{(j)} = \xi_{j_1} \dots \xi_{j_{k_1}}$ and $\eta^{(l)} = \eta_{l_1} \dots \eta_{l_{k_2}}$. A direct calculus shows that

$$\{H_0, \xi^{(j)}\eta^{(l)}\} = -i\Omega(j,l)\xi^{(j)}\eta^{(l)}$$

with

$$\Omega(j,l) := \omega_{j_1} + \dots + \omega_{j_{k_1}} - \omega_{l_1} - \dots - \omega_{l_{k_2}}.$$

Let

$$Q = \sum_{(j,l)\in[1,\dots,n]^k} a_{jl}\xi^{(j)}\eta^{(l)}$$

where $(j,l) \in [1, \dots, n]^k$ means that $j \in [1, \dots, n]^{k_1}$ and $l \in [1, \dots, n]^{k_2}$ with $k_1 + k_2 = k$. Then defining

$$b_{jl} = i\Omega(j,l)^{-1}a_{ij}, \quad c_{jl} = 0 \quad \text{when } \Omega(j,l) \neq 0$$

and

$$c_{jl} = a_{ij}, \quad b_{jl} = 0 \quad \text{when } \Omega(j,l) = 0\,,$$

the polynomials

$$\chi = \sum_{(j,l)\in[1,\dots,n]^k} b_{j,l}\xi^{(j)}\eta^{(l)}$$

and

$$Z = \sum_{(j,l)\in[1,\dots,n]^k} c_{j,l}\xi^{(j)}\eta^{(l)}$$

satisfy (3.7) and (3.8). Furthermore, that Q is real is a consequence of the symmetry relation: $\bar{a}_{jl} = a_{lj}$. Taking into acount that $\Omega_{lj} = -\Omega_{jl}$, this symmetry remains satisfied for the polynomials χ and Z. □

To complete the proof of theorem 3.2, it remains to consider the non resonant case. Recall that we use lemma 3.6 to remove succesively parts of the polynomials P_k for $k = 3, \dots, r$. Therefore the $\Omega_{j,l}$ that we need to consider can be written $k \cdot \omega$ for a $k \in \mathbb{Z}^n$ satisfying $|k| \le r$. Thus if ω is nonresonant up to order r, these $\Omega_{j,l}$ can vanish only if $j = l$ and thus the normal terms constructed in lemma 3.6 have the form $Z = \sum_j a_{j,j}\xi^{(j)}\eta^{(j)} = \sum_j a_{j,j}I^{(j)}$, i.e. Z depends only on the actions. □

Exercise 3.7. — Let $Q = \xi_1\eta_2^2$ (resp. $Q = \xi_1\eta_2^2 + \xi_1^2\eta_2$) and assume $\omega_1/\omega_2 \notin \mathbb{Q}$. Compute the corresponding χ, Z. Then compute the new variables $(\xi', \eta') = \tau^{-1}(\xi, \eta)$, τ being the Lie transform generated by χ. Verify that $H_0(\xi', \eta') = H_0(\xi, \eta) + Q(\xi, \eta)$ (resp. $H_0(\xi', \eta') = H_0(\xi, \eta) + Q(\xi, \eta) +$ order 4).

4. A Birkhoff normal form theorem in infinite dimension

In this section we want to generalize the Birkhoff normal form theorem stated and proved in section 3 in finite dimension to the case of infinite dimension. In view of section 3, we can previse that we face to difficulty: first we have to replace the definition 2.1 by a concept that makes sense in infinite dimension. This will be done in definition 4.4. The second difficulty is more structural: we have to define a class of perturbations P, and in particular a class of polynomials, in which the Birkhoff procedure can apply even with an infinite number of monomials. Concretly, the problem is to be able to verify that at each step of the procedure the formal polynomials that we construct are at least continuous function on the phase space (the continuity of a polynomial is not automatic in infinite dimension since they may contain an infinite number of monomials). This class of polynomial is defined in definition 6.1

and is directly inspired by a class of multilinear forms introduced in [**DS04**], [**DS05**]. In section 4.1 we define our model of infinite dimensional integrable Hamiltonian system and in section 4.2 we state the Birkhoff type result and its dynamical consequences.

4.1. The model. — To begin with we give an abstract model of infinite dimensional Hamiltonian system. In section 5.1 we will give some concrete PDEs that can be described in this abstract framework.

We work in the phase space $\mathcal{P}_s \equiv \mathcal{P}_s(\mathbb{R}) := l^2_s(\mathbb{R}) \times l^2_s(\mathbb{R})$ where, for $s \in \mathbb{R}$, $l^2_s(\mathbb{R}) := \{(a_j)_{j\geq 1} \in \mathbb{R}^{\mathbb{N}} \mid \sum_{j\geq 1} j^{2s}|a_j|^2\}$ is a Hilbert space for the standart norm: $\|a\|^2_s = \sum_{j\geq 1} |j|^{2s}|a_j|^2$.

Let us denote by $(\cdot,\cdot)$ the l^2-scalar product on $l^2_s(\mathbb{R})$. Let $\mathcal{U}$ be an open subset of $l^2_s(\mathbb{R})$, for $F \in C^1(\mathcal{U},\mathbb{R})$ and $a \in \mathcal{U}$, we define the l^2 gradient $\nabla F(a)$ by

$$DF(a)\cdot h = (\nabla F(a), h), \quad \text{for all } h \in l^2_s(\mathbb{R})$$

where $DF(a)$ denotes the differential of F at the point a. We write

$$\nabla F(a) \equiv \left(\frac{\partial F}{\partial a_j}(a)\right)_{j\geq 1}.$$

Notice that, without further hypothesis on F, we only have $\nabla F(a) \in l^2_{-s}(\mathbb{R})$. Then we endow $\mathcal{P}_s$ with the canonical symplectism $\sum_{j\geq 1} dq_j \wedge dp_j$ and we define the Hamiltonian vector field of a regular Hamiltonian function on an open subset $\mathcal{U}$ of $\mathcal{P}_s$, $H \in C^\infty(\mathcal{U},\mathbb{R})$ by

$$X_H(q,p) = \begin{pmatrix} \left(\frac{\partial H}{\partial p_j}(q,p)\right)_{j\geq 1} \\ -\left(\frac{\partial H}{\partial q_j}(q,p)\right)_{j\geq 1} \end{pmatrix}.$$

Again without further hypothesis on H, we only know that $X_H(q,p) \in l^2_{-s}(\mathbb{R}) \times l^2_{-s}(\mathbb{R})$. However, in order to consider the flow of the Hamilton's equations

$$\frac{d}{dt}\begin{pmatrix} q \\ p \end{pmatrix} = X_H(q,p)$$

on $\mathcal{P}_s$, we prefer that the vector field preserves this phase space[(4)], i.e. $X_H(q,p) \in \mathcal{P}_s$ for $(q,p) \in \mathcal{P}_s$. Thus we will be interested in the following class of admissible Hamiltonian functions

[(4)]this condition is not really necessary, $X_H(q,p)$ could be unbounded as an operator from $\mathcal{P}_s$ to $\mathcal{P}_s$

Definition 4.1. — *Let $s \geq 0$, we denote by $\mathcal{H}^s$ the space of real valued functions H defined on a neighborhood $\mathcal{U}$ of the origin in $\mathcal{P}_s$ and satisfying*

$$H \in C^\infty(\mathcal{U}, \mathbb{R}) \quad \textit{and} \quad X_H \in C^\infty(\mathcal{U}, \mathcal{P}_s).$$

In particular the Hamiltonian vector fields of functions F, G in $\mathcal{H}^s$ are in $l_s^2(\mathbb{R}) \times l_s^2(\mathbb{R})$ and we can define their Poisson bracket by

$$\{F, G\}(q,p) = \sum_{j\geq 1} \frac{\partial F}{\partial q_j}(q,p)\frac{\partial G}{\partial p_j}(q,p) - \frac{\partial F}{\partial p_j}(q,p)\frac{\partial G}{\partial q_j}(q,p)\ .$$

We will also use the complex variables

$$\xi_j = \frac{1}{\sqrt{2}}(q_j + ip_j), \quad \eta_j = \frac{1}{\sqrt{2}}(q_j - ip_j).$$

We have $(\xi, \eta) \in \mathcal{P}_s(\mathbb{C})$, the complexification of $\mathcal{P}_s(\mathbb{R})$. In these variables, the Poisson bracket of two functions in $\mathcal{H}^s$ reads

$$\{F, G\} = i\sum_{j\geq 1} \frac{\partial F}{\partial \xi_j}\frac{\partial G}{\partial \eta_j} - \frac{\partial F}{\partial \eta_j}\frac{\partial G}{\partial \xi_j}$$

where

$$\frac{\partial}{\partial \xi_j} = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial q_j} - i\frac{\partial}{\partial p_j}\right), \quad \frac{\partial}{\partial \eta_j} = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial q_j} + i\frac{\partial}{\partial p_j}\right).$$

As in the finite dimensional case, we will say that a function F defined in the variable (ξ, η) is real when $F(\xi, \bar{\xi})$ is real which means that in the original real variables (q,p), F is real valued. We will use the notation $z = (\ldots, \xi_2, \xi_1, \eta_1, \eta_2, \ldots) \in l_s^2(\bar{\mathbb{Z}}, \mathbb{C})$ where $\bar{\mathbb{Z}} = \mathbb{Z} \setminus \{0\}$. We will also denote $\bar{\mathbb{N}} = \mathbb{N} \setminus \{0\}$.

Our model of integrable system is the harmonic oscillator

$$H_0 = \sum_{j\geq 1} \omega_j \xi_j \eta_j$$

where $\omega = (\omega_j)_{j\geq 1} \in \mathbb{R}^{\mathbb{N}}$ is the frequencies vector. We will assume that these frequencies grow at most polynomilally, i.e. that there exist $C > 0$ and $d \geq 0$ such that for any $j \in \bar{\mathbb{N}}$,

$$|\omega_j| \leq C|j|^d, \tag{4.1}$$

in such a way that H_0 be well defined on $\mathcal{P}_s$ for s large enough. The perturbation term is a real function, $P \in \mathcal{H}^s$, having a zero of order at least 3 at the origin. Our Hamiltonian function is then given by

$$H = H_0 + P$$

and the Hamilton's equations read, in the real variables,

$$\left\{\begin{array}{ll} \dot q_j = & \omega_j p_j + \frac{\partial P}{\partial p_j}, \ \ j\geq 1\\ \dot p_j = & -\omega_j q_j - \frac{\partial P}{\partial q_j}, \ j\geq 1\end{array}\right. \tag{4.2}$$

and in the complex ones

$$\left\{\begin{array}{ll} \dot \xi_j = - & i\omega_j \xi_j - i\frac{\partial P}{\partial \eta_j}, \ \ j\geq 1\\ \ \dot \eta_j = & i\omega_j \eta_j + i\frac{\partial P}{\partial \xi_j}, \ j\geq 1.\end{array}\right. \tag{4.3}$$

Our theorem will require essentially two hypotheses: one on the perturbation P (see definition 4.2 below) and one on the frequencies vector ω (see definition 4.4 below). We begin by giving a motivation for these intriguing definitions.

As in the finite dimensional case, the game will consist in removing iteratively, by a canonical transform, the cubic terms of P that are not in normal form with respect to H_0, then the quartic ones and so on. The basic rule remains the same: to remove the monomial $a_j z_{j_1}\ldots z_{j_k}$, we have to control the monomial $\frac{a_j}{\Omega(j)} z_{j_1}\ldots z_{j_k}$ where, as in the finite dimensional case (cf. proof of lemma 3.6), $\Omega(j)=\mathrm{sign}(j_1)\omega_{j_1}+\ldots+\mathrm{sign}(j_k)\omega_j$ is the small divisor. In contrast with the finite dimensional case, the number of monomials that we have to remove at each step is, a priori, infinite. Fortunately, the vector field of many of them are already small in the l^2_s-norm for s large enough:
Consider the simple case where P is a monomial of degree k, $P=z_{j_1}\ldots z_{j_k}$. Assume that $|j_1|\leq\ldots\leq|j_k|$ and that the three largest indexes are large and of the same order, say $N\leq|j_{k-2}|\leq|j_{k-1}|\leq|j_k|\leq 2N$. Then one gets

$$\begin{aligned}\|X_P(z)\|_s^2 &= \sum_{l\in\bar{\mathbb{Z}}} l^{2s}\left|\frac{\partial P}{\partial z_l}(z)\right|^2\\ &\leq \sum_{l=1}^k |j_l|^{2s}|z_{j_1}|^2\ldots|z_{j_{l-1}}|^2|z_{j_{l+1}}|^2\ldots|z_{j_k}|^2\\ &\leq \sum_{l=1}^k \frac{|j_l|^{2s}}{\prod_{m\neq l}|j_m|^{2s}}\prod_{m\neq l}|j_m|^{2s}|z_{j_m}|^2\\ &\leq \frac{C}{N^{2s}}\|z\|_s^{2k-2}\end{aligned}$$

which is small for large N. This calculus explains why we will control only the small divisors that involve at most two large indexes, whence the definition 4.4 below. Concerning the class of nonliniearities P that we can consider, the preceding calculus does not suffice to justify the precise form of definition

4.2 below but again it explains why the estimate (4.4) involves only the three largest indexes. Actually some other constraints are required like the control of the l^2_s-norm of X_P (cf. proposition 6.1) and the stability of the class under Poisson bracketing (cf. proposition 6.3).

For $j \in \bar{\mathbb{Z}}^k$ with $k \geq 3$, we define $\mu(j)$ as the third largest integer between $|j_1|, \ldots, |j_k|$. Then we set $S(j) := |j_{i_0}| - |j_{i_1}| + \mu(j)$ where $|j_{i_0}|$ and $|j_{i_1}|$ are respectively the largest integer and the second largest integer between $|j_1|, \ldots, |j_k|$. In particular

$$\text{if } |j_1| \leq \ldots \leq |j_k| \text{ then } \mu(j) := |j_{k-2}| \text{ and } S(j) = |j_k| - |j_{k-1}| + |j_{k-2}|.$$

For $j \in \bar{\mathbb{Z}}^k$ with $k \leq 2$, we fix $\mu(j) = S(j) = 1$.

Definition 4.2. — *Let $k \geq 3$, $N \in \mathbb{N}$ and $\nu \in [0, +\infty)$ and let*

$$Q(\xi, \eta) \equiv Q(z) = \sum_{l=0}^{k} \sum_{j \in \bar{\mathbb{Z}}^l} a_j z_{j_1} \ldots z_{j_l}$$

be a formal polynomial of degree k on $\mathcal{P}_s(\mathbb{C})$. Q is in the class $\mathcal{T}_k^{N,\nu}$ if there exists a constant $C > 0$ such that for all j

$$|a_j| \leq C \frac{\mu(j)^{N+\nu}}{S(j)^N}. \tag{4.4}$$

We will see in section 6.1 that $\mathcal{T}_k^{N,\nu} \subset \mathcal{H}^s$ for $s \geq \nu + 1/2$ (cf. proposition 6.1) and thus in particular a polynomial in $\mathcal{T}_k^{N,\nu}$ is well defined on a neighborhood of the origin in $\mathcal{P}_s(\mathbb{C})$ for s large enough. The best constant C in (4.4) defines a norm for which $\mathcal{T}_k^{N,\nu}$ is a Banach space. We set

$$\mathcal{T}_k^{\infty,\nu} = \cap_{N \in \mathbb{N}} \mathcal{T}_k^{N,\nu}$$

and

$$\mathcal{T}^\nu = \cup_{k \geq 0} \mathcal{T}_k^{\infty,\nu}.$$

This definition is similar to a class of multilinear forms first introduced by Delort and Szeftel in [**DS04**] and [**DS05**].

Definition 4.3. — *A function P is in the class $\mathcal{T}$ if*

- *there exist $s_0 \geq 0$ such that, for any $s \geq s_0$, $P \in \mathcal{H}^s$*
- *for each $k \geq 1$ there exits $\nu \geq 0$ such that the Taylor's expansion of degree k of P at zero belongs to $\mathcal{T}_k^{\infty,\nu}$.*

In section 6.1 we will establish some properties of polynomials in $\mathcal{T}_k^{N,\nu}$, in particular we will see that this class has a good behaviour regarding to the Poisson bracket (cf. proposition 6.3).

Concerning the frequencies, we define:

Definition 4.4. — *A frequencies vector $\omega \in \mathbb{R}^{\bar{\mathbb{N}}}$ is* ***strongly non resonant*** *if for any $r \in \bar{\mathbb{N}}$, there are $\gamma > 0$ and $\alpha > 0$ such that for any $j \in \bar{\mathbb{N}}^r$ and any $1 \leq i \leq r$, one has*

$$\left|\omega_{j_1} + \cdots + \omega_{j_i} - \omega_{j_{i+1}} - \cdots - \omega_{j_r}\right| \geq \frac{\gamma}{\mu(j)^\alpha} \tag{4.5}$$

except if $\{j_1, \ldots, j_i\} = \{j_{i+1}, \ldots, j_r\}$.

This definition was first introduced in [**Bam03**].

Remark 4.5. — The direct generalization of the definition 2.1 to the infinite dimensional case would read $\sum_{j \geq 1} \omega_j k_j \neq 0$ for all $k \in \mathbb{Z}^{\bar{\mathbb{N}}} \setminus 0$. But in the infinite dimensional case this condition no more implies that there exists $C(r)$ such that

$$\left|\sum_{j \geq 1} \omega_j k_j\right| \geq C(r) \text{ for all } k \in \mathbb{Z}^{\bar{\mathbb{N}}} \setminus 0 \text{ satisfying } \sum_{j \geq 1} |k_j| \leq r \tag{4.6}$$

which is the property that we used in the proof of theorem 3.2. Actually this last property (4.6) is too restrictive in infinite dimension. For instance when the frequencies are the eigenvalues of the 1-d Schrödinger operator with Dirichlet boundary conditions (cf. example 5.5), one shows that $\omega_j = j^2 + a_j$ where $(a_j)_{j \geq 1} \in l^2$ (cf. for instance [**Mar86, PT87**]) and thus if l is an odd integer and $j = (l^2 - 1)/2$ then one has $\omega_{j+1} - \omega_j - \omega_l \to_{l \to \infty} 0$.

Our strongly nonresonant condition says that $|\sum_{j \geq 1} \omega_j k_j|$ is controled from below by a quantity which goes to zero when the third largest index of the frequencies involved grows to infinity, but the length of k is fixed. Precisely one has:

Proposition 4.6. — *A frequencies vector $\omega \in \mathbb{R}^{\bar{\mathbb{N}}}$ is strongly non resonant if and only if for any $r \in \bar{\mathbb{N}}$, there are $\gamma > 0$ and $\alpha > 0$ such that for any $N \in \bar{\mathbb{N}}$*

$$|\sum_{m=1}^{N} \omega_m k_m + k_{l_1} \omega_{l_1} + k_{l_2} \omega_{l_2}| \geq \frac{\gamma}{N^\alpha} \tag{4.7}$$

for any indexes $l_1, l_2 > N$ *and for any* $k \in \bar{\mathbb{Z}}^{N+2} \setminus \{0\}$ *with* $\sum_{m=1}^{N} |k_m| \leq r$, $|k_{l_1}| + |k_{l_2}| \leq 2$.

In this form, the strongly nonresonant condition may be compared to the so called Melnikov condition used in the KAM theory (cf. section 7.2).

Proof. — In order to see that the first form implies the second ones, we remark that the expression $\sum_{m=1}^{N} \omega_m k_m + k_{l_1}\omega_{l_1} + k_{l_2}\omega_{l_2}$ may be rewrite as $\omega_{j_1} + \cdots + \omega_{j_i} - \omega_{j_{i+1}} - \cdots - \omega_{j_{r'}}$ for some $r \leq r' \leq r+2$ and for some $j \in \bar{\mathbb{N}}^{r'}$ satisfying $\mu(j) \leq N$.
Conversely, $\omega_{j_1} + \cdots + \omega_{j_i} - \omega_{j_{i+1}} - \cdots - \omega_{j_r}$ may be rewrite as $\sum_{m=1}^{N} \omega_m k_m + k_{l_1}\omega_{l_1} + k_{l_2}\omega_{l_2}$ with $N = \mu(j)$, $k_{l_1}, k_{l_2} = \pm 1$ and $\sum_{m=1}^{N} |k_m| \leq r - 2$. □

4.2. The statement. — We can now state our principal result:

Theorem 4.7. — *Assume that* P *belongs to the class* $\mathcal{T}$ *and that* ω *is strongly non resonant and satisfies* (4.1) *for some* $d \geq 0$. *Then for any* $r \geq 3$ *there exists* s_0 *and for any* $s \geq s_0$ *there exists* $\mathcal{U}_s$, $\mathcal{V}_s$ *neighborhoods of the origin in* $\mathcal{P}_s(\mathbb{R})$ *and* $\tau_s : \mathcal{V}_s \to \mathcal{U}_s$ *a real analytic canonical transformation which is the restriction to* $\mathcal{V}_s$ *of* $\tau := \tau_{s_0}$ *and which puts* $H = H_0 + P$ *in normal form up to order* r *i.e.*

$$H \circ \tau = H_0 + Z + R$$

with

(i) Z *is a continuous polynomial of degree* r *with a regular vector field (i.e.* $Z \in \mathcal{H}^t$ *for all* $t \geq 0$*) which only depends on the actions:* $Z = Z(I)$.
(ii) $R \in \mathcal{H}^s(\mathcal{V}_s, \mathbb{R})$ *and* $\|X_R(q,p)\|_s \leq C_s \|(q,p)\|_s^r$ *for all* $(q,p) \in \mathcal{V}_s$.
(iii) τ *is close to the identity:* $\|\tau(q,p) - (q,p)\|_s \leq C_s \|(q,p)\|_s^2$ *for all* $(q,p) \in \mathcal{V}_s$.

This theorem was first proved in [**BG04**] under a slightly more general hypothesis on the perturbation (cf. remark 6.2 and 6.6).

This theorem says, as in the finite dimensional case, that we can change the coordinates in a neighborhood of the origin in such a way that the Hamiltonian be integrable up to order r, r being fixed at the principle. Remark that the concept of integrability that we gave in definition 2.4 does not directly extend to the infinite dimensional case(5). However, if a Hamiltonian $H(q,p)$ depends only on the actions I_j, $j \geq 1$ then we can say that H is integrable in the sense

(5)It can be done with an appropriate definition of linear independence of an infinity of vector fields.

that we can integrate it. Actually the solutions to the Hamilton's equation in the variables (I,θ) are given by

$$\begin{cases} \theta_j(t) = \theta_j(0) + t\omega_j, & j \geq 1 \\ I_j(t) = I_j(0), & j \geq 1. \end{cases}$$

The proof of theorem 4.7, that we will present in section 6, is very closed to the proof of theorem 2.6 in [**BDGS05**]. The dynamical consequences of this theorem are similar as those of the Birkhoff theorem in finite dimension:

Corollary 4.8. — *Assume that P belongs to the class $\mathcal{T}$ and that ω is strongly non resonant. For each $r \geq 3$ and $s \geq s_0(r)$, there exists $\varepsilon_0 > 0$ and $C > 0$ such that if $\|(q_0,p_0)\| = \varepsilon < \varepsilon_0$ the solution $(q(t),p(t))$ of the Hamiltonian system associated to H which takes value (q_0,p_0) at $t = 0$ sastisfies*

(i)
$$\|(q_t,p_t)\| \leq 2\varepsilon \quad \text{for } |t| \leq \frac{C}{\varepsilon^{r-1}}.$$

(ii) *for each $j = 1,\ldots,n$*
$$|I_j(t) - I_j(0)| \leq \frac{\varepsilon^3}{|j|^{2s}} \quad \text{for } |t| \leq \frac{C}{\varepsilon^{r-2}}$$

(iii) *let $r_1 + r_2 = r+1$, then there exists a torus $\mathcal{T}_0 \subset \mathcal{P}_s$ such that for each $s' < s-1$,*
$$dist_{s'}((q(t),p(t),\mathcal{T}_0) \leq C_{s'}\varepsilon^{r_1} \quad \text{for } |t| \leq 1/\varepsilon^{r_2}$$
where $dist_s$ denotes the distance on $\mathcal{P}_s$ associated with the norm $\|\cdot\|_s$

Remark 4.9. — This corollary remains valid for *any small initial datum* and this makes a big difference with the dynamical consequences of the KAM-type result where one has to assume that the initial datum belongs to a Cantor-type set (cf. section 7.2). But of course the result is not the same, here the stability is guaranteed only for long, *but finite*, time. When KAM theory applies, the stability is inconditional, i.e. guaranteed for infinite time. Furthermore, the KAM theorem does not require that the perturbation be in the class $\mathcal{T}$.

Remark 4.10. — The first assertion implies in particular that the Hamiltonian system have *almost global solutions*: if the initial datum is smaller than ϵ then the solution exits and is controled (in the initial norm) for times of order ϵ^{-r}, the order r being arbitrarily fixed at the principle. This consequence can be very interesting in the context of PDEs for which the global existence is not known (cf. [**BDGS05**]).

This result was first proved in [**BG04**]. For convenience of the reader we repeat it here.

Proof. — The proof is similar to the proof of corollary 3.3, we focus on the slight differences. Denote $z := (q, p)$ and $z' = \tau^{-1}(z)$ where τ is the transformation given by theorem 4.7. Then we define $N(z) := \|z\|_s^2 = 2\sum_{j=1}^{\infty} j^{2s} I_j(q, p)$. Using that Z depends only on the actions, we have

$$\{N, H\}(z) = \{N \circ \tau, H \circ \tau\} \circ \tau^{-1}(z) = \{N \circ \tau, R\}(z').$$

Therefore, as in the finite dimensional case, we get $|\dot{N}| \leq CN^{(r+1)/2}$ and assertions (i) and (ii) follow.

To prove (iii), denote by $\bar{I}_j := I_j(0)$ the initial actions in the normalized coordinates. Up to the considered times

$$\left|I_j(t) - \bar{I}_j\right| \leq \frac{C\epsilon^{2r_1}}{j^{2s}} \ . \tag{4.8}$$

Then, as in the proof of corollary 3.4, we define the torus

$$\mathcal{T}_0 := \left\{z \in \mathcal{P}_s \ : \ I_j(z) = \bar{I}_j \ , j \geq 1\right\}.$$

We have for $s' < s - 1$

$$d_{s'}(z(t), \mathcal{T}_0) \leq \left[\sum_j j^{2s'} \left|\sqrt{I_j(t)} - \sqrt{\bar{I}_j}\right|^2\right]^{1/2} . \tag{4.9}$$

Notice that for $a, b \geq 0$,

$$\left|\sqrt{a} - \sqrt{b}\right| \leq \sqrt{|a - b|} \ .$$

Thus, using (4.8), we obtain

$$[d_{s'}(z(t), \mathcal{T}_0)]^2 \leq \sum_j \frac{j^{2s}|I_j(t) - \bar{I}_j|}{j^{2(s-s')}} \leq \sup_j \left(j^{2s}|I_j(t) - \bar{I}_j|\right) \sum \frac{1}{j^{2(s-s')}}$$

which is convergent provided $s' < s - 1/2$. □

5. Application to Hamiltonian PDEs

In this section we first descibe two concrete PDE's and we then verify that the abstract results of section 4.2 apply to them.

5.1. Examples of 1-d Hamiltonian PDEs. — Two examples of 1-d Hamiltonian PDEs are given: the nonlinear wave equation and the nonlinear Schrödinger equation. In [**Cra00**], the reader may find much more examples like the Korteweg-de Vries equation, the Fermi-Pasta-Ulam system or the waterwaves system.
In section 7.1 we will comment on recent generalisation to some d-dimensional PDE with $d \geq 2$.

Nonlinear wave equation. — As a first concrete example we consider a 1-d nonlinear wave equation

$$u_{tt} - u_{xx} + V(x)u = g(x,u) \ , \quad x \in S^1 \ , \ t \in \mathbb{R} \ , \tag{5.1}$$

with Dirichlet boundary condition: $u(0,t) = u(\pi,t) = 0$ for any t. Here V is a 2π periodic C^∞ non negative potential and $g \in C^\infty(S^1 \times \mathcal{U})$, $\mathcal{U}$ being a neighbourhood of the origin in $\mathbb{R}$. For compatibility reasons with the Dirichlet conditions, we further assume that $g(x,u) = -g(-x,-u)$ and that V is even. Finally we assume that g has a zero of order two at $u = 0$ in such a way that $g(x,u)$ appears, in the neighborhood of $u = 0$, as a perturbation term.

Defining $v = u_t$, (5.1) reads

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ u_{xx} - V(x)u + g(x,u) \end{pmatrix} .$$

Furthermore, let $H : H^1(S^1) \times L^2(S^1) \mapsto \mathbb{R}$ defined by

$$H(u,v) = \int_{S^1} \left(\frac{1}{2}v^2 + \frac{1}{2}u_x^2 + \frac{1}{2}Vu^2 + G(x,u) \right) dx \tag{5.2}$$

where G is such that $\partial_u G = -g$, then (5.1) reads as an Hamiltonian system

$$\begin{aligned} \partial_t \begin{pmatrix} u \\ v \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -u_{xx} + Vu + \partial_u G \\ v \end{pmatrix} \\ &= J\nabla_{u,v} H(u,v) \end{aligned} \tag{5.3}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ represents the symplectic structure and where $\nabla_{u,v} = \begin{pmatrix} \nabla_v \\ \nabla_v \end{pmatrix}$ with ∇_u and ∇_v denoting the L^2 gradient with respect to u and v respectively.
Define the operator $A := (-\partial_{xx} + V)^{1/2}$, and introduce the variables (p,q) given by

$$q := A^{1/2}u \ , \quad p := A^{-1/2}v \ .$$

Then, on $H^s(S^1)\times H^s(S^1)$ with $s\geq 1/2$, the Hamiltonian (5.2) takes the form H_0+P with

$$H_0(q,p)=\frac{1}{2}\left(\langle Ap,p\rangle_{L^2}+\langle Aq,q\rangle_{L^2}\right)$$

and

$$P(q,p)=\int_{S^1}G(x,A^{-1/2}q)dx$$

Now denote by $(\omega_j)_{j\in\bar{\mathbb{N}}}$ the eigenvalues of A with Dirichlet boundary conditions and ϕ_j, $j\in\bar{\mathbb{N}}$, the associated eigenfunctions, i.e.

$$A\phi_j=\omega_j\phi_j.$$

For instance, for $V=0$, we have $\phi_j(x)=\sin jx$ and $\omega_j=j$.
An element (q,p) of $H^s(S^1)\times H^s(S^1)$ satisfying the Dirichlet boundary conditions may be decomposed on the Hilbert basis $(\phi_j)_{j\geq 1}$:

$$q(x)=\sum_j q_j\phi_j(x)\quad\text{and}\quad p(x)=\sum_j p_j\phi_j(x)$$

with $(q_j,p_j)_{j\geq 1}\in\mathcal{P}_s=l^2_s(\mathbb{R})\times l^2_s(\mathbb{R})$. Then the Hamiltonian of the non linear wave equation (5.1) reads on $\mathcal{P}_s$

$$H=\sum_{j\geq 1}\omega_j\frac{p_j^2+q_j^2}{2}+P$$

where P belongs in $C^\infty(\mathcal{P}_s,\mathbb{R})$ and has a zero of order at least 3 at the origin and $\mathcal{P}_s$ is endowed with the same symplectic structure as in section 4.1, i.e. the Hamilton's equations read as in (4.2).

Nonlinear Schrödinger equation. — As a second example we consider the nonlinear Schrödinger equation

$$-\mathrm{i}\dot{\psi}=-\psi_{xx}+V\psi+\partial_3 g(x,\psi,\bar{\psi})\ ,\quad x\in S^1,\quad t\in\mathbb{R}\tag{5.4}$$

with Dirichlet boundary conditions: $\psi(0,t)=\psi(\pi,t)=0$ for any t. Here V is a 2π periodic C^∞ potential. We assume that $g(x,z_1,z_2)$ is $C^\infty(S^1\times\mathcal{U})$, $\mathcal{U}$ being a neighbourhood of the origin in $\mathbb{C}\times\mathbb{C}$. The notation ∂_3 means that we take the partial derivative with respect to the third argument. We also assume that g has a zero of order three at $(z_1,z_2)=(0,0)$ and that $g(x,z,\bar{z})\in\mathbb{R}$. To deal with Dirichlet boundary conditions we have to ensure the invariance of the phase space under the vector field associated with the equation, to this end we assume that V is even and that $g(-x,-z,-\bar{z})=g(x,z,\bar{z})$.

Defining the Hamiltonian function of the system as

$$H=\int_{S^1}\frac{1}{2}\left(|\psi_x|^2+V|\psi|^2\right)+g(x,\psi(x),\bar{\psi}(x))dx,\tag{5.5}$$

equation (5.4) is equivalent to

$$\dot{\psi} = i\nabla_{\bar{\psi}} H$$

where i represents a symplectic structure.
Let A be the Sturm–Liouville operator $-\partial_{xx} + V$ with Dirichlet boundary conditions, the frequencies ω_j, $j \geq 1$, are the corresponding eigenvalues and the normal modes ϕ_j are the corresponding eigenfunctions. We can write $H = H_0 + P$ with, for $(\psi, \bar{\psi}) \in H^1(S^1) \times H^1(S^1)$,

$$H_0(\psi, \bar{\psi}) = \langle A\psi, \bar{\psi} \rangle_{L^2}$$

and

$$P(\psi, \bar{\psi}) = \int_{S^1} g(x, \psi(x), \bar{\psi}(x)) dx.$$

As in the previous example an element $(\psi, \bar{\psi})$ of $H^s(S^1) \times H^s(S^1)$ satisfying the Dirichlet boundary conditions may be decomposed on the Hilbert basis $(\phi_j)_{j \geq 1}$:

$$\psi(x) = \sum_j \xi_j \phi_j(x) \quad \text{and} \quad \bar{\psi}(x) = \sum_j \eta_j \phi_j(x)$$

with $(\xi_j, \eta_j)_{j\geq 1} \in \mathcal{P}_s(\mathbb{C}) = l^2_s(\mathbb{C}) \times l^2_s(\mathbb{C})$. Then the Hamiltonian of the non linear Schrödinger equation (5.4) reads on $\mathcal{P}_s(\mathbb{C})$

$$H = \sum_{j\geq 1} \omega_j \xi_j \eta_j + P.$$

Here P belongs to $C^\infty(\mathcal{P}_s, \mathbb{C})$, satisfies $P(u, \bar{u}) \in \mathbb{R}$ and has a zero of order at least 3 at the origin. On the other hand $\mathcal{P}_s(\mathbb{C})$ is endowed with the same symplectic structure as in section 4.1, i.e. the Hamilton's equations read as in (4.3).

Notice that defining p and q as the real and imaginary parts of ψ, namely write $\psi = p + \mathrm{i}q$ we can recover the real form (4.2).

5.2. Verification of the hypothesis. — The dynamical consequences of our Birkhoff normal form theorem for PDEs are given in corollary 4.8, in particular the solution remains under control in the H^s-norm during a very long time if the H^s-norm of the initial datum is small. But this suppose that the Hamiltonian function of the PDE satisfies the two conditions: strong non resonancy of the linear frequencies and perturbation term in the good class.

Verification of the condition on the perturbation term. — We work in the general framework of 1-d PDEs given in section 4.1. The Hamiltonian perturbation reads

$$P(q,p) = \int_{S^1} f(x, q(x), p(x))dx \tag{5.6}$$

where $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$, $q(x) = \sum_{j\geq 1} q_j\phi_j(x)$, $p(x) = \sum_{j\geq 1} p_j\phi_j(x)$ and $((q_j)_{j\geq 1}, (p_j)_{j\geq 1}) \in \mathcal{P}_s$. Here $(\phi_j)_{j\geq 1}$ are the eigenfunctions of the selfadjoint operator A and form a basis of the phase space. That P belongs to the class $\mathcal{T}^\nu$ is directly in relation with the distribution of the ϕ_j's. Actually we have

Proposition 5.1. — *Let $\nu \geq 0$. Assume that for each $k \geq 1$ and for each $N \geq 0$ there exists $C > 0$ such that for all $j \in \mathbb{N}^k$*

$$\left|\int_{S^1} \phi_{j_1} \dots \phi_{j_k} dx\right| \leq C\frac{\mu(j)^{N+\nu}}{S(j)^N} \tag{5.7}$$

then any P of the general form (5.6) *satisfying the symmetries imposed by the domain of the operator A is in the class $\mathcal{T}^\nu$.*

Proof. — The Taylor's polynomial of P at order n reads

$$P_n = \sum_{k=0}^{n} \sum_{(j,l)\in\bar{\mathbb{N}}^{k_1}\times\bar{\mathbb{N}}^{k_2}} a_{jl} q_{j_1} \dots q_{j_{k_1}} p_{l_1} \dots p_{l_{k_2}}$$

with

$$a_{jl} = \frac{1}{k_1!k_2!}\int_{S^1} \frac{\partial^k f}{\partial_2^{k_1}\partial_3^{k_2}}(x,0,0)\phi_{j_1}(x)\dots\phi_{j_{k_1}}(x)\phi_{l_1}(x)\dots\phi_{l_{k_2}}(x).$$

Since P satisfies the symmetry conditions imposed by the domain of A, we can decompose $\frac{\partial^k f}{\partial_2^{k_1}\partial_3^{k_2}}(x,0,0)$ on the Hilbert basis $(\phi_m)_{m\geq 1}$:

$$\frac{\partial^k f}{\partial_2^{k_1}\partial_3^{k_2}}(x,0,0) = \sum_{m\in\bar{\mathbb{N}}} b_m\phi_m(x)\,.$$

Thus we get

$$a_{jl} = \frac{1}{k_1!k_2!}\sum_{m\in\bar{\mathbb{N}}} b_m c_{mjl}$$

where

$$c_{mjl} = \int_{S^1} \phi_m(x)\phi_{j_1}(x)\dots\phi_{j_{k_1}}(x)\phi_{l_1}(x)\dots\phi_{l_{k_2}}(x).$$

By hypothesis

$$|c_{mjl}| \leq C\mu(m,j,l)^{N+\nu}S(m,j,l)^{-N}$$

where $\mu(m,j,l)=\mu(m,j_1,\dots,j_{k_1},l_1,\dots,l_{k_2})$ and $S(m,j,l)=S(m,j_1,\dots,j_{k_1},l_1,\dots,l_{k_2})$. So it remains to verify that there exists $C>0$ such that

$$\sum_{m\in\bar{\mathbb{N}}}|b_m|\mu(m,j,l)^{N+\nu}S(m,j,l)^{-N}\leq C\mu(j,l)^{N+\nu}S(j,l)^{-N}\ . \tag{5.8}$$

This last inequality is a consequence of the following facts:

- For each i there exists C_i such that $|b_m|(1+m)^i\leq C_i$ for all $m\geq 1$ (because f is infinitely smooth and the b_m act as Fourier coefficient relative to the basis (ϕ_j)).
- If $m\leq\mu(j,l)$ then $\mu(m,j,l)=\mu(j,l)$ and $S(m,j,l)=S(j,l)$.
- If $m>\mu(j,l)$ then $\mu(m,j,l)<m$ and thus $\sum(1+m)^{-i}\mu(m,j,l)^{N+\nu}$ converges for i large enough.
- If $m>\mu(j,l)$ then $(1+m)S(m,j,l)\geq S(j,l)$ (exercise).

□

So it remains to verify condition (5.7) in concrete cases. We begin with a very simple one:

Lemma 5.2. — *If $\phi_j=e^{ijx}$, $j\in\mathbb{Z}$ then* (5.7) *holds true with* $\nu=0$.

Proof. — We have $\int_{S^1}\phi_{j_1}\dots\phi_{j_k}dx=2\pi$ if $j_1+\ldots+j_k=0$ and $\int_{S^1}\phi_{j_1}\dots\phi_{j_k}dx=0$ if $j_1+\ldots+j_k\neq 0$. So we have to prove that there exists $C>0$ such that for any $j\in\mathbb{Z}^k$ satisfying $j_1+\ldots+j_k=0$,

$$S(j)\leq C\mu(j).$$

By symmetry we can assume that j is ordered, i.e. $|j_1|\leq|j_2|\leq\ldots\leq|j_k|$. In this case, recall that $S(j)=||j_k|-|j_{k-1}||+\mu(j)$. But since $j_1+\ldots+j_k=0$

$$||j_k|-|j_{k-1}||\leq|j_k+j_{k-1}|\leq\sum_{m=1}^{k-2}|j_m|\leq(k-2)\mu(j).$$

Therefore $S(j)\leq(k-1)\mu(j)$. □

The condition (5.7) can be verified in a much more general case:

Definition 5.3. — *A sequence of functions $(\phi_j)_{j\geq 1}$ is **well localised with respect to the exponentials** if, writing $\phi_j=\sum_{l\in\mathbb{Z}}\phi_j^l e^{ilx}$, for each $n\geq 0$ there exists $c_n>0$ such that*

$$|\phi_j^l|\leq\frac{c_n}{\min_\pm(1+|l\pm j|)^n} \tag{5.9}$$

for all $j,l\in\mathbb{Z}$.

Example 5.4. — If $A = -\partial_{xx}$ with Dirichlet boundary conditions, then $\phi_j(x) = \sin jx$ for $j \geq 1$ which are well localized with respect to the exponentials.

Example 5.5. — Let $A = -\partial_{xx} + V$ with Dirichlet boundary conditions, where V is a C^∞, 2π periodic potential. Then $\phi_j(x)$ are the eigenfunctions of a Sturm Liouville operator. By the theory of Sturm Liouville operators (cf [**Mar86, PT87**]) they are well localized with respect to the exponentials (cf [**CW93**]).

This last example applies to both PDE's we have considered in section 5.1.

Proposition 5.6. — *If $(\phi_j)_{j\geq 1}$ is well localised with respect to the exponentials then the condition* (5.7) *holds true with $\nu = 0$.*

Proof. — For a multi-index $l \in \mathbb{Z}^k$, we denote $[l] = l_1 + \ldots + l_k$.
Assume that (5.9) is satisfied then

$$\left|\int_{S^1} \phi_{j_1} \ldots \phi_{j_k} dx\right| = 2\pi \left|\sum_{l\in\mathbb{Z}^k,[l]=0} \phi_{j_1}^{l_1} \ldots \phi_{j_k}^{l_k}\right|$$

$$\leq c_n^k \sum_{l\in\mathbb{Z}^k,[l]=0} \prod_{i=1}^{k} \frac{1}{\min_\pm(1+|l_i \pm j_i|)^n}. \tag{5.10}$$

On the other hand, define $\varepsilon_i = \pm 1$ in such a way that $|l_i + \varepsilon_i j_i| = \min_\pm |l_i \pm j_i|$. By symmetry we can restrict our analysis to the case where j is ordered: $|j_1| \leq \ldots \leq |j_k|$. Then if $\varepsilon_k\varepsilon_{k-1} = -1$ we write using $[l] = 0$,

$$|j_k - j_{k-1}| = |\varepsilon_k j_k - l_k + \varepsilon_{k-1} j_{k-1} - l_{k-1} + \ldots + \varepsilon_1 j_1 - l_1 - \sum_{i=1}^{k-2} \varepsilon_i j_i|,$$

to conclude

$$|j_k - j_{k-1}| \leq (k-2)\mu(j) + D(l,j)$$

where

$$D(l,j) = \sum_{i=1}^{k} |l_i - \varepsilon_i j_i|.$$

If $\varepsilon_k\varepsilon_{k-1} = 1$ we obtain similarly

$$|j_k + j_{k-1}| \leq (k-2)\mu(j) + D(l,j).$$

Hence, since $S(j) = \mu(j) + |j_k| - |j_{k-1}| = \mu(j) + \min_\pm |j_k \pm j_{k-1}|$, we obtain in both cases

$$S(j) \leq (k-1)\mu(j) + D(l,j).$$

As a consequence we have

$$\frac{\mu(j)}{S(j)} \geq \frac{1}{k-1} \frac{1}{1+D(l,j)} \ . \tag{5.11}$$

Finally notice that, by definition of ϵ_i,

$$\prod_{i=1}^{k} \min_{\pm}(1+|l_i \pm j_i|) \geq 1+D(l,j).$$

Inserting this last inequality and (5.11) in (5.10) leads to

$$\begin{aligned}\left|\int_{S^1} \phi_{j_1}\dots\phi_{j_k} dx\right| &\leq 2\pi c_n^k \sum_{l\in\mathbb{Z}^k,[l]=0} \frac{1}{(1+D(l,j))^n} \\ &\leq 2\pi (k-1)^N c_n^k \frac{\mu(j)^N}{S(j)^N} \sum_{l\in\mathbb{Z}^k,[l]=0} \frac{1}{(1+D(l,j))^{n-N}} \ .\end{aligned}$$

The last sum converges for $n > N+k-1$ and thus (5.7) is verified. □

Verification of the strong non resonancy condition in a simple case. — This subsection is inspired by section 5 in [**BG04**], actually the case considered here is much more simple.
Let A be the operator on $L^2(-\pi,\pi)$ defined by

$$Au = -\frac{d^2u}{dx^2} + V \star u$$

where V is a 2π periodic potential and $\star$ denotes the convolution product:

$$V \star u(x) = \int_{-\pi}^{\pi} V(x-y)u(y)dy \ .$$

We consider A with Dirichlet boundary conditions, i.e. on the domain $D(A)$ of odd and 2π-periodic H^2 function (cf. section 5.1),

$$D(A) = \{u(x) = \sum_{j\geq 1} u_j \sin jx \mid (u_j)_{j\geq 1} \in l_2^2(\mathbb{N},\mathbb{R})\}.$$

We assume that V belongs to the following space ($m \geq 1$)

$$V \in \mathcal{W}_m := \{V(x) = \frac{1}{\pi}\sum_{j\geq 1} \frac{v_j}{(1+|j|)^m} \cos jx \mid v_j \in [-1/2, 1/2],\ j\geq 1\}$$

that we endow with the product probability measure. Notice that a potential in $\mathcal{W}_m$ is in the Sobolev space H^{m-1} and that we assume V even to leave invariant $D(A)$ under the convolution product by V.

In this context the frequencies are given by

$$\omega_j = j^2 + \frac{v_j}{(1+|j|)^m} \ , \quad j \geq 1$$

and one has

Theorem 5.7. — *There exists a set $F_m \subset \mathcal{W}_m$ whose measure equals 1 such that if $V \in F_m$ then the frequencies vector $(\omega_j)_{j\geq 1}$ is strongly non resonnant.*

Remark 5.8. — A similar result holds true when considering the more interesting case $A = -\frac{d^2}{dx^2} + V$ which corresponds to our non linear Schrödinger equation (5.4). But the proof is much more complicated (cf. [**BG04**]).
In the case of the non linear wave equation (5.1) with a constant potential $V = m$, the frequencies reads $\omega_j = \sqrt{j^2+m}$ and it is not too difficult to prove that these frequencies satisfy (4.5) for most choices of m (see [**Bam03**] or [**DS04**]).

Instead of proving theorem 5.7, we prefer to prove the following equivalent statement

Proposition 5.9. — *Fix $r \geq 1$ and $\gamma > 0$ small enough. There exist positive constants $C \equiv C_r$, $\alpha \equiv \alpha(r,\gamma)$, $\delta \equiv \delta(r,\gamma) \leq \gamma$ and a set $F_{r,\gamma} \subset \mathcal{W}_m$ whose measure is larger than $1 - C\gamma$ such that if $V \in F_{r,\gamma}$ then for any $N \geq 1$*

$$\left| \sum_{j=1}^{N} k_j\omega_j + \epsilon_1\omega_{l_1} + \epsilon_2\omega_{l_2} \right| \geq \frac{\delta}{N^\alpha} \tag{5.12}$$

for any $k \in \mathbb{Z}^N$ with $|k| := \sum_{j=1}^N |k_j| \leq r$, for any indexes $l_1,\ l_2 > N$, and for any $\epsilon_1, \epsilon_2 \in \{0, 1, -1\}$ except if
$k = 0$ and $\epsilon_1 = \epsilon_2 = 0$.

Theorem 5.7 is deduced from proposition 5.9 by defining

$$F_m := \cap_{r\geq 1} \cup_{\gamma>0} F_{r,\gamma} \ ,$$

and remarking that this is numerable intersection of sets with full measure.

In order to prove proposition 5.9, we first prove that $\sum_{j=1}^N k_j\omega_j$ cannot accumulate on $\mathbb{Z}$. Precisely we have

Lemma 5.10. — *Fix $r \geq 1$ and $\gamma > 0$ small enough. There exist positive constants $C \equiv C_r$, $\beta \equiv \beta(r,\gamma)$ and a set $F'_{r,\gamma} \subset \mathcal{W}_m$ whose measure equals*

$1-C\gamma$ such that if $V\in F'_{r,\gamma}$ then for any $N\geq 1$ and any $b\in\mathbb{Z}$

$$|\sum_{j=1}^{N}k_j\omega_j-b|\geq\frac{\gamma}{N^\beta}\ , \tag{5.13}$$

for any $k\in\mathbb{Z}^N$ with $0<|k|\leq r$.

Proof First notice that, given $(a_1,\dots,a_r)\neq 0$ in $\mathbb{Z}^r$, $M>0$ and $c\in\mathbb{R}$ the Lesbegue measure of

$$\{x\in[-M,M]^r\mid|\sum_{i=1}^{r}a_ix_i+c|<\delta\}$$

is smaller than $(2M)^{r-1}\delta$. Hence given $k\in\mathbb{Z}^N$ of length less than r and $b\in\mathbb{Z}$ the Lesbegue measure of

$$\mathcal{X}_k:=\left\{x\in[-1/2,1/2]^N\ :\ \left|\sum_{j=1}^{N}k_j(j^2+x_j)+b\right|<\frac{\gamma}{N^\beta}\right\}$$

is smaller than γ/N^β. Now consider the set

$$\left\{v\in[-1/2,1/2]^N\ :\ \left|\sum_{j=1}^{N}k_j\omega_j-b\right|<\frac{\gamma}{N^\beta}\right\},$$

it is contained in the set of the v's such that $(v_j/(1+|j|)^m)\in\mathcal{X}_k$. The measure of this set in turn is estimated by $(1+N)^{m(r-1)}\gamma/N^\beta$. To conclude the proof we have to sum over all the k's and the b's. To count the cardinality of the set of the k's and the b's to be considered remark that if $|\sum_{j=1}^N k_j\omega_j-b|\geq\delta$ with $\delta<1$ then $|b|\leq 1+|\sum_{j=1}^N k_j\omega_j|\leq 1+(1+N^2)r$. So that to guarantee (5.13) for all possible choices of k, b and N, it suffices to remove from $\mathcal{W}_m$ a set of measure

$$\sum_{N\geq 1}\gamma\frac{1}{N^\beta}(1+N)^{m(r-1)}N^r(1+(1+N^2)r)\ .$$

Choosing $\beta:=r(1+m)+4$, the last series converges and the lemma is proved. □

Proof of proposition 5.9 First of all, for $\epsilon_1=\epsilon_2=0$, (5.12) is a direct consequence of lemma 5.10 choosing $\alpha=\beta$, $\delta=\gamma$ and $F_{r,\gamma}=F'_{r,\gamma}$.

When $\epsilon_1 = \pm 1$ and $\epsilon_2 = 0$, (5.12) reads

$$|\sum_{j=1}^{N} k_j\omega_j \pm \omega_l| \geq \frac{\delta}{N^\alpha} \tag{5.14}$$

for some $l \geq N$. Notice that $|\sum_{j=1}^{N} k_j\omega_j| \leq r(N^2+1)$ and thus, if $l > 2Nr$, (5.14) is always true. When $l \leq 2Nr$, we apply lemma 5.10 replacing r by $r+1$ and N by $2Nr$ to obtain (5.14) with $\alpha = \beta(r+1,\gamma)$, $\delta = \frac{\gamma}{(2r)^\alpha}$ and $F_{r,\gamma} = F'_{r+1,\gamma}$. In the same way one proves (5.12) when $\epsilon_1\epsilon_2 = 1$. So it remains to establish an estimate of the form

$$|\sum_{j=1}^{N} k_j\omega_j + \omega_{l_1} - \omega_{l_2}| \geq \gamma\frac{\delta}{N^\alpha} \tag{5.15}$$

for any $k \in \mathbb{Z}^N$, $0 < |k| \leq r$ and for any $N \leq l_1 \leq l_2$.
One has

$$\omega_{l_1} - \omega_{l_2} = l_1^2 - l_2^2 + \frac{v_{l_1}}{(1+|l_1|)^m} - \frac{v_{l_2}}{(1+|l_2|)^m} \ .$$

Therefore if $4^{\frac{1}{m}} N^{\frac{\beta+1}{m}} \gamma^{\frac{-1}{m}} \leq l_1 \leq l_2$, one has with $b = l_1^2 - l_2^2$

$$|\omega_{l_1} - \omega_{l_2} - b| \leq \frac{\gamma}{2N^\alpha} \ .$$

Thus using lemma 5.10, (5.15) holds true with $\alpha = \beta + 1$, $\delta = \gamma/2$ and for $F_{r,\gamma} = F'_{r,\gamma}$.
Finally assume $l_1 \leq 4^{\frac{1}{m}} N^{\frac{\beta+1}{m}} \gamma^{\frac{-1}{m}}$, taking into acount $|\sum_{j=1}^{N} k_j\omega_j| \leq r(N^2+1)$, (5.15) is satisfied when $l_2 \geq 4^{\frac{1}{m}} N^{\frac{\beta+1}{m}} \gamma^{\frac{-1}{m}} 3r$. So it remains to consider the case when $l_1 \leq l_2 \leq 12rN^{\frac{\beta+1}{m}} \gamma^{\frac{-1}{m}}$. But in this case, we can apply lemma 5.10 with r replaced by $r+2$ and N replaced by $12rN^{\frac{\beta+1}{m}} \gamma^{\frac{-1}{m}}$ to obtain (5.15) with $\alpha = \frac{\beta(r+2,\gamma)(\beta(r+2,\gamma)+1)}{m}$, $\delta = \gamma(12r\gamma^{\frac{-1}{m}})^{-\beta(r+2,\gamma)}$ and $F_{r,\gamma} = F'_{r+2,\gamma}$. □

6. Proof of our Birkhoff theorem in infinite dimension

We first have to study the class of polynomials that we introduce in section 4.2.

6.1. Preliminary results on polynomials in $\mathcal{T}_k^{N,\nu}$. — The two propositions given in this section were first proved, in a different context, in [**DS04**], [**DS05**]. Nevertheless, for convenience of the reader, we present slightly different proofs in our context.

Proposition 6.1. — *Let $k \in \bar{\mathbb{N}}$, $N \in \mathbb{N}$, $\nu \in [0, +\infty)$, $s \in \mathbb{R}$ with $s > \nu + 3/2$, and let $P \in \mathcal{T}^{N,\nu}_{k+1}$. Then*

(i) *P extends as a continuous polynomials on $\mathcal{P}_s(\mathbb{C})$ and there exists a constant $C > 0$ such that for all $z \in \mathcal{P}_s(\mathbb{C})$*

$$|P(z)| \leq C \, \|z\|_s^{k+1}$$

ii) *Assume moreover that $N > s + 1$, then the Hamiltonian vector field X_P extends as a bounded function from $\mathcal{P}_s(\mathbb{C})$ to $\mathcal{P}_s(\mathbb{C})$. Furthermore, for any $s_0 \in (\nu + 1, s]$, there is $C > 0$ such that for any $z \in \mathcal{P}_s(\mathbb{C})$*

$$\|X_P(z)\|_s \leq C \, \|z\|_s \, \|z\|_{s_0}^{(k-1)} \, . \tag{6.1}$$

Remark 6.2. — The estimate (6.1) is of tame type (see [**AG91**] for a general presentation of this concept) and has to be compared with the classical tame estimate

$$\|uv\|_{H^s} \leq C_s(\|u\|_{H^s} \, \|v\|_{H^1} + \|u\|_{H^1} \, \|v\|_{H^s}) \quad \forall u, v \in H^s(\mathbb{R}).$$

On the other hand, in [**BG04**], we obtained a Birkhoff normal form theorem for perturbations whose Taylor's polynomials satisfies a more general tame estimate. In this sense the theorem obtained there is more general.

Proof. — (i) Without loss of generality we can assume that P is an homogeneous polynomial of degree $k+1$ in $\mathcal{T}^{N,\nu}_{k+1}$ and we write for $z \in \mathcal{P}_s(\mathbb{C})$

$$P(z) = \sum_{j \in \bar{\mathbb{Z}}^{k+1}} a_j \, z_{j_1} \ldots z_{j_{k+1}} \, . \tag{6.2}$$

One has, using first (4.4) and then $\frac{\mu(j)}{S(j)} \leq 1$,

$$\begin{aligned}
|P(z)| &\leq C \sum_{j \in \bar{\mathbb{Z}}^{k+1}} \frac{\mu(j)^{N+\nu}}{S(j)^N} \prod_{i=1}^{k+1} |z_{j_i}| \\
&\leq C \sum_{j \in \bar{\mathbb{Z}}^{k+1}} \frac{\mu(j)^{\nu}}{\prod_{i=1}^{k+1} |j_i|^s} \prod_{i=1}^{k+1} |j_i|^s |z_{j_i}| \\
&\leq C \sum_{j \in \bar{\mathbb{Z}}^{k+1}} \frac{1}{\prod_{i=1}^{k+1} |j_i|^{s-\nu}} \prod_{i=1}^{k+1} |j_i|^s |z_{j_i}| \\
&\leq C \left(\sum_{l \in \bar{\mathbb{Z}}} \frac{1}{|l|^{2s-2\nu}} \right)^{\frac{k+1}{2}} \|z\|_s^{k+1}
\end{aligned}$$

where in the last inequality we used $k+1$ times the Cauchy-Schwarz inequality. Since $s > \nu + 1/2$, the last sum converges and the first assertion is proved.

(ii) The Hamiltonian vector field of the polynomial (6.2) reads $X_P(z) = (\frac{\partial P}{\partial z_l}(z))_{l\in\bar{\mathbb{Z}}}$ with, for $l \in \bar{\mathbb{Z}}$

$$\frac{\partial P}{\partial z_l}(z) = \sum_{j\in\bar{\mathbb{Z}}^{k+1}} \sum_{i=1}^{k+1} \delta_{j_i,l}\ a_{j_1\ldots j_{i-1}lj_{i+1}\ldots j_{k+1}}\ z_{j_1}\ldots z_{j_{i-1}}z_{j_{i+1}}\ldots z_{j_{k+1}},$$

where $\delta_{m,n}$ denotes the Kronecker symbol. Since the estimate (4.4) is symmetric with respect to $j_1,\ldots,j_{k+1}$ we deduce

$$\left|\frac{\partial P}{\partial z_l}(z)\right| \leq C\ (k+1) \sum_{j\in\bar{\mathbb{Z}}^k} \frac{\mu(j,l)^{N+\nu}}{S(j,l)^N}|z_{j_1}|\ldots|z_{j_k}|$$

where $\mu(j,l) = \mu(j_1,\ldots,j_k,l)$ and $S(j,l) = S(j_1,\ldots,j_k,l)$. So we obtain

$$\|X_P(z)\|_s^2 = \sum_{l\in\bar{\mathbb{Z}}} l^{2s}\left|\frac{\partial P}{\partial z_l}(z)\right|^2$$

$$\leq C(k+1)^2 \sum_{l\in\bar{\mathbb{Z}}}\Big(\sum_{j\in\bar{\mathbb{Z}}^k}\frac{l^s\mu(j,l)^{N+\nu}}{S(j,l)^N}|z_{j_1}|\ldots|z_{j_k}|\Big)^2. \tag{6.3}$$

By symmetry we may restrict ourselves to ordered multi-indices j, i.e. satisfying $|j_1| \leq \ldots \leq |j_k|$. We then notice that for all $l \in \bar{\mathbb{Z}}$ and for all ordered $j \in \bar{\mathbb{Z}}^k$

$$l\frac{\mu(j,l)}{S(j,l)} \leq 2|j_k|. \tag{6.4}$$

Actually if $|l| \leq 2|j_k|$ then (6.4) holds true since $\frac{\mu(j,l)}{S(j,l)} \leq 1$. Now if $|l| \geq 2|j_k|$ then $S(l,j) \geq ||l| - |j_k|| \geq 1/2|l|$ and thus

$$l\frac{\mu(j,l)}{S(j,l)} \leq 2\mu(j,l) \leq 2|j_k|$$

since j is ordered.

Fix $\epsilon > 0$ such that $N - s \geq 1+\epsilon$ and $2s_0 \geq 2\nu + 3 + \epsilon$. Inserting (6.4) in (6.3) and using $\mu(j,l) \leq S(j,l)$ we get (here C designs a generic constant depending on k, N, ν, s and s_0)

$$\|X_P(z)\|_s^2 \leq C\sum_{l\in\bar{\mathbb{Z}}}\Big(\sum_{j\in\bar{\mathbb{Z}}^k_>}\frac{|j_k|^s\ \mu(j,l)^{\nu+1+\epsilon}}{S(j,l)^{1+\epsilon}}|z_{j_1}|\ldots|z_{j_k}|\Big)^2.$$

where $\bar{\mathbb{Z}}^k_>$ denotes ther space of ordered multi-indices.
Now we use that, for ordered j, $\mu(j,l) \le |j_{k-1}|$ and[6] $S(j,l) \ge 1 + |l - j_k|$ to obtain

$$\begin{aligned}
\|X_P(z)\|_s^2 &\le C \sum_{l\in\bar{\mathbb{Z}}} \Big(\sum_{j\in\bar{\mathbb{Z}}^k_>} \frac{|j_k|^s \ |j_{k-1}|^{\nu+1+\epsilon}}{(1+|l-j_k|)^{1+\epsilon}} |z_{j_1}| \ldots |z_{j_k}| \Big)^2 \\
&= C \sum_{l\in\bar{\mathbb{Z}}} \Big(\sum_{j_k\in\bar{\mathbb{Z}}} A_{j_k} \ B_{j_k} \Big)^2
\end{aligned}$$

where

$$\begin{aligned}
A_{j_k} &= |j_k|^s |z_{j_k}|, \\
B_{j_k} &= \sum_{(j_1,\ldots,j_{k-1})\in\Delta_{j_k}} \frac{|j_{k-1}|^{\nu+1+\epsilon}}{(1+|l-j_k|)^{1+\epsilon}} \prod_{i=1}^{k-1} |z_{j_i}|
\end{aligned}$$

and $\Delta_{j_k} = \{(j_1,\ldots,j_{k-1}) \in \bar{\mathbb{Z}}^{k-1}_> \mid j_{k-1} \le j_k\}$. Therefore using the Cauchy-Schwarz inequality we get

$$\|X_P(z)\|_s^2 \le C \, \|z\|_s^2 \sum_{l\in\bar{\mathbb{Z}}} \sum_{j_k\in\bar{\mathbb{Z}}} \frac{1}{(1+|l-j_k|)^{2+2\epsilon}} \Big(\sum_{(j_1,\ldots,j_{k-1})\in\Delta_{j_k}} \prod_{i=1}^{k-1} \alpha_{j_i} \ \beta_{j_i} \Big)^2$$

where, for $i = 1, \ldots, k-1$,

$$\alpha_{j_i} = |j_i|^{s_0} |z_{j_i}|,$$

and

$$\begin{aligned}
\beta_{j_i} &= \frac{1}{|j_i|^{s_0}}, \quad \text{for } i = 1, \ldots, k-2, \\
\beta_{j_{k-1}} &= \frac{1}{|j_{k-1}|^{s_0-\nu-1-\epsilon}}.
\end{aligned}$$

Then, applying $k-1$ times the Cauchy-Schwarz, we obtain (6.1). □

The second essential property of polynomials in $\mathcal{T}_k^{N,\nu}$ is captured in the following

Proposition 6.3. — *The map $(P,Q) \mapsto \{P,Q\}$ define a continuous map from $\mathcal{T}_{k_1+1}^{N,\nu_1} \times \mathcal{T}_{k_2+1}^{N,\nu_2}$ to $\mathcal{T}_{k_1+k_2}^{N',\nu'}$ for any $N' < N - \max(\nu_1,\nu_2) - 1$ and any $\nu' > \nu_1 + \nu_2 + 1$.*

[6] for $l \ge j_{k-1}$, $S(j,l) = |j_{k-1}| + |l - j_k|$ and for $l \le j_{k-1}$, $S(j,l) \ge l$

Proof. — As in the proof of proposition 6.1, we assume that $P \in \mathcal{T}^{N,\nu_1}_{k_1+1}$ and $Q \in \mathcal{T}^{N,\nu_2}_{k_2+1}$ are homogeneous polynomial and we write

$$P(z) = \sum_{j \in \bar{\mathbb{Z}}^{k_1+1}} a_j \, z_{j_1} \dots z_{j_{k_1+1}}$$

and

$$Q(z) = \sum_{i \in \bar{\mathbb{Z}}^{k_2+1}} b_i \, z_{i_1} \dots z_{i_{k_2+1}} \ .$$

In view of the symmetry of the estimate (4.4) with respect to the involved indices, one easily obtains

$$\{P, Q\}(z) = \sum_{(j,i) \in \bar{\mathbb{Z}}^{k_1+k_2}} c_{j,i} \, z_{j_1} \dots z_{j_{k_1}} z_{i_1} \dots z_{i_{k_2}}$$

whith

$$|c_{j,i}| \leq (k_1+1)(k_2+1) \sum_{l \in \bar{\mathbb{Z}}} \frac{\mu(j,l)^{N+\nu_1}}{S(j,l)^N} \frac{\mu(i,l)^{N+\nu_2}}{S(i,l)^N}.$$

Therefore it remains to prove that there exists $C > 0$ such that for all $j \in \bar{\mathbb{Z}}^{k_1}$ and all $i \in \bar{\mathbb{Z}}^{k_2}$,

$$\sum_{l \in \bar{\mathbb{Z}}} \frac{\mu(j,l)^{N+\nu_1}}{S(j,l)^N} \frac{\mu(i,l)^{N+\nu_2}}{S(i,l)^N} \leq C \frac{\mu(j,i)^{N'+\nu'}}{S(j,i)^{N'}} \tag{6.5}$$

In order to simplify the notation, and because it does not change the estimates of (6.5), we will assume $k_1 = k_2 = k$. We can also assume by symmetry that

- all the indices are positive: $j_1, \dots, j_k, i_1, \dots, i_k \geq 1$.
- j and i are ordered: $j_1 \leq \dots \leq j_k$ and $i_1 \leq \dots \leq i_k$.
- $j_k \geq i_k$.

Then we consider two cases: $i_k \geq j_{k-1}$ and $i_k \leq j_{k-1}$.

First case: $\mathbf{j_k \geq i_k \geq j_{k-1}}$
We first remark that in this case,

$$\mu(j,l) \leq \mu(i,j) \text{ and } \mu(i,l) \leq \mu(i,j). \tag{6.6}$$

For any multi-index j we denote $\tilde{S}(j) = S(j) - \mu(j)$, i.e. $S(j)$ is the difference between the two largest indices. We have for all $l \in \bar{\mathbb{Z}}$

$$\tilde{S}(i,j) \leq \tilde{S}(i,l) + \tilde{S}(j,l). \tag{6.7}$$

Actually, if $|l| \leq i_k$ then $\tilde{S}(j,l) = |j_k - \max(j_{k-1}, |l|)| \geq j_k - i_k = \tilde{S}(i,j)$, if $|l| \geq j_k$ then $\tilde{S}(i,l) = |l| - i_k \geq j_k - i_k = \tilde{S}(i,j)$ and if $j_k \leq |l| \leq i_k$ then

$\tilde{S}(i,l)+\tilde{S}(j,l)=j_k-i_k=\tilde{S}(i,j)$.
Combining 6.6 and 6.7 we get

$$\frac{\mu(i,j)}{S(i,j)}\geq 1/2\ \min\left(\frac{\mu(i,l)}{S(i,l)},\frac{\mu(j,l)}{S(j,l)}\right).$$

Assume for instance that $\frac{\mu(j,l)}{S(j,l)}\leq\frac{\mu(i,l)}{S(i,l)}$ and let $\varepsilon>0$. We then have

$$\begin{aligned}\sum_{l\in\bar{\mathbb{Z}}}\frac{\mu(j,l)^{N+\nu_1}}{S(j,l)^N}\frac{\mu(i,l)^{N+\nu_2}}{S(i,l)^N}&\leq 2^{N-1-\epsilon}\sum_{l\in\bar{\mathbb{Z}}}\mu(j,l)^{\nu_1}\frac{\mu(i,j)^{N-1-\varepsilon}}{S(i,j)^{N-1-\varepsilon}}\frac{\mu(i,l)^{1+\varepsilon+\nu_2}}{S(i,l)^{1+\varepsilon}}\\&\leq C\frac{\mu(j,i)^{N+\nu_1+\nu_2}}{S(j,i)^{N-1-\varepsilon}}\end{aligned}$$

where we used (6.6) and the fact that $\sum_{l\in\bar{\mathbb{Z}}}\frac{1}{S(i,l)^{1+\epsilon}}\leq C$ uniformly with respect to i. Hence in this first case, (6.5) holds true with $N'=N-1-\varepsilon$ and $\nu'=\nu_1+\nu_2+1+\varepsilon$.

Second case: $\mathbf{j_k\geq j_{k-1}\geq i_k}$

In this second case (6.7) remains true. Actually if $|l|\leq j_{k-1}$ then $\tilde{S}(i,j)=\tilde{S}(j)=\tilde{S}(j,l)$ and if $|l|\geq j_{k-1}$ then $\tilde{S}(j,l)+\tilde{S}(i,l)=|j_k-|l||+|i_k-|l||\geq j_k-j_{k-1}=\tilde{S}(i,j)$. Unfortunately (6.6) may be wrong. More precisely, we still have $\mu(i,j)\geq\mu(i,l)$ but it may occur that $\mu(i,j)<\mu(j,l)$. Now, if we further assume that $\mu(j,l)\leq 2\mu(i,j)$, we can proceed as in the first case to obtain the same estimates with an irrelevant extra factor $2^{N+\nu_1}$. So it remains to consider indices l for which $\mu(j,l)>2\mu(i,j)$. Notice that it can occur only if $|l|\geq j_{k-2}$, and thus we have $\mu(j,l)\leq|l|$. Further, as $\mu(i,j)\geq i_k$, we deduce $|l|\geq 2i_k$ and thus $|i_k-l|\geq l/2$.
We finally have to argument differently depending wether $\tilde{S}(i,l)\leq\tilde{S}(j,l)$ or not. If it is true then, in view of (6.7), we get

$$S(i,j)\leq\mu(i,j)+2\tilde{S}(j,l)\leq 2S(j,l).$$

Thus, using that $\mu(j,l)\leq l$, $\mu(i,l)\leq\mu(i,j)$ and $|i_k-l|\geq l/2$,

$$\begin{aligned}&\sum_{l,\ \tilde{S}(i,l)\leq\tilde{S}(j,l)}\frac{\mu(j,l)^{N+\nu_1}}{S(j,l)^N}\frac{\mu(i,l)^{N+\nu_2}}{S(i,l)^N}\\&\qquad\leq 2^{N-1-\epsilon-\nu_1}\sum_{l\in\bar{\mathbb{Z}}}\frac{l^N}{S(i,j)^{N-1-\varepsilon-\nu_1}}\frac{\mu(i,j)^{N+\nu_2}}{(|i_k-l|+i_{k-1})^N}\frac{1}{S(j,l)^{1+\varepsilon}}\\&\qquad\leq C\frac{\mu(j,i)^{N+\nu_2}}{S(j,i)^{N-1-\varepsilon-\nu_1}}\end{aligned}$$

where, as usual, we used that $\sum_{l\in\bar{\mathbb{Z}}}\frac{1}{S(j,l)^{1+\epsilon}}\le C$ uniformly with respect to j. It remains to consider the subcase $\tilde{S}(i,l)\ge\tilde{S}(j,l)$. We then have

$$\frac{S(i,j)}{\mu(i,j)}\le 2\frac{S(i,l)}{\mu(i,l)}$$

and thus, using again $\mu(j,l)\le l$, $\mu(i,l)\le\mu(i,j)$ and $|i_k-l|\ge l/2$,

$$\begin{aligned}\sum_{l,\ \tilde{S}(i,l)\ge\tilde{S}(j,l)} & \frac{\mu(j,l)^{N+\nu_1}}{S(j,l)^N}\frac{\mu(i,l)^{N+\nu_2}}{S(i,l)^N}\\ &\le 2^{N-1-\epsilon-\nu_1}\sum_{l\in\bar{\mathbb{Z}}}\frac{l^{\nu_1}}{(|i_k-l|+i_{k-1})^{\nu_1}}\frac{\mu(i,j)^{N-1-\varepsilon-\nu_1}}{S(i,j)^{N-1-\varepsilon-\nu_1}}\frac{\mu(i,l)^{\nu_1+\nu_2+1+\varepsilon}}{S(i,l)^{1+\varepsilon}}\\ &\le C\frac{\mu(j,i)^{N+\nu_2}}{S(j,i)^{N-1-\varepsilon-\nu_1}}\end{aligned}$$

Hence, in the second case, (6.5) holds true with $N'=N-1-\varepsilon-\nu_1$ and $\nu'=\nu_1+\nu_2+1+\varepsilon$. □

We end this section with a corollary concerning Lie transforms associated to polynomials in $\mathcal{T}^\beta$.

Corollary 6.4. — *Let χ be a real homogeneous polynomial in $\mathcal{T}_l^{\infty,\beta}$ with $\beta\ge 0$, $l\ge 3$ and denote by ϕ the associated Lie transform.*
(i) Let $F\in\mathcal{H}^s$ with s large enough, then $F\circ\phi\in\mathcal{H}^s$.
(ii) Let $P\in\mathcal{T}_n^{\infty,\nu}$, $\nu\ge 0$, $n\ge 3$ and fix $r\ge n$ an integer. Then

$$P\circ\phi=Q_r+R_r$$

where:
- Q_r is a polynomial of degree r belonging to $\mathcal{T}_r^{\infty,\nu'}$ with
$\nu'=\nu+(r-n)(\beta+1)+2$,
- R_r is a real Hamiltonian in the class $\mathcal{T}$ having a zero of order $r+1$ at the origin.

Proof. — (i) Let $\mathcal{W}$ be a neighborhood of 0 in $\mathcal{P}_s$ such that F belongs to $\mathcal{H}^s(\mathcal{W},\mathbb{R})$. Since $\chi\in\mathcal{T}_l^{\infty,\beta}$, by proposition 6.1, $\chi\in\mathcal{H}^s$ for $s>s_1=\beta+3/2$. In particular, for $s>s_1$, the flow Φ^t generated by the vector field X_χ transports an open subset of $\mathcal{P}_s$ into an open subset of $\mathcal{P}_s$. Furthermore, since χ has a zero of order 3, there exists $\mathcal{U}$ a neighborhood of 0 in $\mathcal{P}_s$ such that the flow $\mathcal{U}\ni(q,p)\mapsto\Phi^t(q,p)\in\mathcal{W}$ is well defined and smooth for $0\le t\le 1$. By definition of the Lie transform, $\phi=\Phi^1$. In view of the formula

$$X_{F\circ\phi}(q,p)=(D\phi(q,p))^{-1}X_F(\phi(q,p)),$$

we deduce that $F\circ\phi\in\mathcal{H}^s$ for $s>s_1$.
(ii) We use lemma 2.8 (which remains valid in infinite dimension) to conclude

$$\frac{d^k}{dt^k}P\circ\phi^t(q,p)\big|_{t=0}=P^{(k)}(q,p)$$

where $P^{(k+1)}=\{P^{(k)},\chi\}$ and $P^{(0)}=P$. Therefore applying the Taylor's formula to $P\circ\phi^t(q,p)$ between $t=0$ and $t=1$ we deduce

$$P\circ\phi(q,p)=\sum_{k=0}^{r-n}\frac{1}{n!}P^{(k)}(q,p)+\frac{1}{(r-n)!}\int_0^1(1-t)^rP^{(r-n+1)}(\Phi^t(q,p))dt. \tag{6.8}$$

Notice that $P^{(k)}(q,p)$ is a homogeneous polynomial of degree $n+k(l-2)$ and, by propositions 6.1 and 6.3, $P^{(k)}(q,p)\in\mathcal{T}^{\nu+k\beta+k+2}\cap\mathcal{H}^s$ for $s\geq\nu+k\beta+k+2$. Therefore (6.8) decomposes in the sum of a polynomial of degree r in $\mathcal{T}_r^{\infty,\nu'}$ and a function in $\mathcal{H}^s$ having a zero of degree $r+1$ at the origin.

□

6.2. Proof of theorem 4.7. — We are now in position to prove theorem 4.7. Actually the proof is very close to the proof of theorem 3.2, i.e. the finite dimensional case. So again having fixed some $r\geq 3$, the idea is to construct iteratively for $k=2,\dots,r$, a neighborhood $\mathcal{V}_k$ of 0 in $\mathcal{P}_s$ (s large enough depending on r), a canonical transformation τ_k, defined on $\mathcal{V}_k$, an increasing sequence $(\nu_k)_{k=2,\dots,r}$ of positive numbers and real Hamiltonians Z_k,P_{k+1},Q_{k+2},R_k such that

$$H_k:=H\circ\tau_k=H_0+Z_k+P_{k+1}+Q_{k+2}+R_k \tag{6.9}$$

and with the following properties

(i) Z_k is a polynomial of degree k in $\mathcal{T}_k^{\infty,\nu_k}$ having a zero of order 3 at the origin and Z_k depends only on the (new) actions: $\{Z_k,I_j\}=0$ for all $j\geq 1$.
(ii) P_{k+1} is a homogeneous polynomial of degree $k+1$ in $\mathcal{T}_{k+1}^{\infty,\nu_k}$.
(iii) Q_{k+2} is a polynomial of degree $r+1$ in $\mathcal{T}_{r+1}^{\infty,\nu_k}$ having a zero of order $k+2$ at the origin.
(iv) R_k is a regular Hamiltonian belonging to $\mathcal{H}^s(\mathcal{V}_k,\mathbb{R})$ for s large enough and having a zero of order $r+2$ at the origin.

First we fix $s>\nu_r+3/2$ to be sure to be able to apply proposition 6.1 at each step (ν_r will be defined later on independently of s). Then we notice that (6.9) at order r proves theorem 4.7 with $Z=Z_r$ and $R=P_{r+1}+R_r$ (notice that

$Q_{r+2}=0$). In particular, by proposition 6.1 assertion (ii), $X_{P_{r+1}}$ satisfies

$$\left\|X_{P_{r+1}}(q,p)\right\|_s \leq C_s \left\|(q,p)\right\|_s^r . \tag{6.10}$$

On the other hand, since R_r belongs to $\mathcal{H}^s$, we can apply the Taylor's formula at order $r+1$ to X_{R_r} to obtain the same estimate (6.10) for R_r on $\mathcal{V}\subset\mathcal{V}_r$ a neighborhood of 0 in $\mathcal{P}_s$.

The Hamiltonian $H=H_0+P$ has the form (6.9) for $k=2$ with $\tau_2=I$, $\nu_2=\nu$, $Z_2=0$, P_3 being the Taylor's polynomial of P of degree 3, Q_4 being the Taylor's polynomial of P of degree $r+1$ minus P_3 and $R_2=P-P_3-Q_4$. We show now how to pass from k to $k+1$.

We search for τ_{k+1} of the form $\tau_k\circ\phi_{k+1}$, ϕ_{k+1} being a Lie transform associated to the Hamiltonian function $\chi_{k+1}\in\mathcal{T}_{k+1}^{\infty,\nu_k'}$ where ν_k' will be determined in lemma 6.5. This Lie transform is well defined and smooth on a neighborhood $\mathcal{V}_{k+1}\subset\mathcal{V}_k$. Recall that by Taylor's formula we get for regular F

$$F\circ\phi_{k+1}=F+\{F,\chi_{k+1}\}+1/2\{\{F,\chi_{k+1}\},\chi_{k+1}\}+\dots$$

We decompose $H_k\circ\phi_{k+1}$ as follows

$$H_k\circ\phi_{k+1} = H_0+Z_k+\{H_0,\chi_{k+1}\}+P_{k+1} \tag{6.11}$$
$$+\ H_0\circ\phi_{k+1}-H_0-\{H_0,\chi_{k+1}\}\ +\ Q_{k+2}\circ\phi_{k+1} \tag{6.12}$$
$$+\ R_k\circ\phi_{k+1}\ +\ Z_k\circ\phi_{k+1}-Z_k\ +\ P_{k+1}\circ\phi_{k+1}-P_{k+1}\ . \tag{6.13}$$

From corollary 6.4 and formula (6.8), we deduce that (6.12) and (6.13) are regular Hamiltonians having a zero of order $k+2$ at the origin and that the sum of these terms decomposes in $P_{k+2}+Q_{k+3}+R_{k+1}$ with P_{k+2}, Q_{k+3} and R_{k+1} satisfying the properties (ii), (iii) and (iv) at rank $k+1$ (with $\nu_{k+1}=k\nu_k'+\nu_k+k+2$). So it remains to prove that χ_{k+1} can be choosen in such way that $Z_{k+1}:=Z_k+\{H_0,\chi_{k+1}\}+P_{k+1}$ satisfies (i). This is a consequence of the following lemma

Lemma 6.5. — *Let $\nu\in[0,+\infty)$ and assume that the frequencies vector of H_0 is strongly non resonant. Let Q be a homogeneous real polynomial of degree k in $\mathcal{T}_k^{\infty,\nu}$, there exist $\nu'>\nu$, homogeneous real polynomials χ and Z of degree k in $\mathcal{T}_k^{\infty,\nu'}$ such that*

$$\{H_0,\chi\}+Q=Z \tag{6.14}$$

and

$$\{Z,I_j\}=0 \quad \forall j\geq 1. \tag{6.15}$$

Proof. — For $j \in \bar{\mathbb{N}}^{k_1}$ and $l \in \bar{\mathbb{N}}^{k_2}$ with $k_1 + k_2 = k$ we denote

$$\xi^{(j)}\eta^{(l)} = \xi_{j_1} \dots \xi_{j_{k_1}} \eta_{l_1} \dots \eta_{l_{k_2}}.$$

One has

$$\{H_0, \xi^{(j)}\eta^{(l)}\} = -i\Omega(j,l)\xi^{(j)}\eta^{(l)}$$

with

$$\Omega(j,l) := \omega_{j_1} + \ldots + \omega_{j_{k_1}} - \omega_{l_1} - \ldots - \omega_{l_{k_2}}.$$

Let $Q \in \mathcal{T}_k^{\infty,\nu}$

$$Q = \sum_{(j,l) \in \bar{\mathbb{N}}^k} a_{jl} \xi^{(j)}\eta^{(l)}$$

where $(j,l) \in \bar{\mathbb{N}}^k$ means that $j \in \bar{\mathbb{N}}^{k_1}$ and $l \in \bar{\mathbb{N}}^{k_2}$ with $k_1 + k_2 = k$. Let us define

$$b_{jl} = i\Omega(j,l)^{-1} a_{ij}, \quad c_{jl} = 0 \quad \text{when } \{j_1, \dots, j_{k_1}\} \neq \{l_1, \dots, l_{k_2}\}$$

and

$$c_{jl} = a_{ij}, \quad b_{jl} = 0 \quad \text{when } \{j_1, \dots, j_{k_1}\} = \{l_1, \dots, l_{k_2}\}.$$

As ω is strongly non resonant, there exist γ and α such that

$$|\Omega(j,l)| \geq \frac{\gamma}{\mu(j,l)^\alpha}$$

for all $(j,l) \in \bar{\mathbb{N}}^k$ with $\{j_1, \dots, j_{k_1}\} \neq \{l_1, \dots, l_{k_2}\}$. Thus, in view of definition 4.2, the polynomials

$$\chi = \sum_{(j,l) \in \bar{\mathbb{N}}^k} b_{j,l} \xi^{(j)}\eta^{(l)},$$

and

$$Z = \sum_{(j,l) \in \bar{\mathbb{N}}^k} c_{j,l} \xi^{(j)}\eta^{(l)}$$

belong in $\mathcal{T}_k^{\infty,\nu'}$ with $\nu' = \nu + \alpha$. Furthermore by construction they satisfy (6.14) and (6.15). Finally, as in the finite dimensional case, that Q is real is equivalent to the symmetry relation: $\bar{a}_{jl} = a_{lj}$. Taking into acount that $\Omega_{lj} = -\Omega_{jl}$, this symmetry remains satisfied for the polynomials χ and Z. □

Remark 6.6. — In this context, when we solve the so-called homological equation (6.14), we loose some regularity ($\nu' = \nu + \alpha$ where α can be very large when r grows). This make a big difference with [**BG04**] where the tame modulus property (and a truncation in the Fourier modes) allowed to solve the homological equation in a fix space (but with growing norm).

7. Generalisations and comparison with KAM type results

7.1. Generalisations of theorem 4.7. — In order to apply theorem 4.7, the main difficulty consists in verifying the strong nonresonancy condition (cf definition 4.4). When we consider 1-d PDE with Dirichlet boundary conditions, this condition is mostly satisfied (see remark 5.8). But, in a lot of other physical situations, the condition (4.5) is too restrictive. Let us describe two examples:

Periodic boudary conditions. — Consider, as in section 5.1, the non linear Schrödinger equation (5.4) but instead of Dirichlet boundary conditions, we now impose periodic boundary conditions: $\psi(x+2\pi,t)=\psi(x,t)$ for all x and t in $\mathbb{R}$. The frequencies are then the eigenvalues of the sturm Liouville operator $A=-\partial_{xx}+V$ with periodic boundary conditions. It turns out (see for instance [**Mar86**]) that these eigenvalues can be indexed by $\bar{\mathbb{Z}}$ in such a way that $\omega_j=j^2+o(1)$, $j\in\bar{\mathbb{Z}}$. In particular we get $\omega_j-\omega_{-j}=o(1)$ and thus (4.5) cannot be satisfied. The same problems appears with the non linear wave equation (5.1) with periodic boundary conditions. However we notice that in both cases the condition (5.9) remains satisfied for the eigenfunctions $(\phi_j)_{j\in\bar{\mathbb{Z}}}$ (see [**CW93**]). That means that nonlinear terms of type (5.6) remains in the class $\mathcal{T}^0$.

Space dimension greater than 2. — Let us descibe the case of the semilinear Klein-Gordon equation on a sphere. Let S^{d-1} be the unit sphere in $\mathbb{R}^d$ $(d\geq 2)$ and Δ_g be the Laplace-Beltrami operator on S^{d-1} for its canonical metric. We consider the nonlinear Klein-Gordon equation

$$(\partial_t^2-\Delta_g+m)v=-\partial_2 f(x,v) \tag{7.1}$$

where m is a strictly positive constant and $f\in C^\infty(S^{d-1}\times\mathbb{R})$ vanishes at least at order 3 in v, $\partial_2 f$ being the derivative with respect to the second variable. The frequencies of the unperturbed problems are the square roots of the eigenvalues of the operator $-\Delta_g+m$:

$$\omega_j=\sqrt{j(j+d-2)+m}\,,\quad j\geq 0.$$

The problem here is that these eigenvalues are no more simple. Denoting by E_j the eigenspace associated to $\lambda_j=j(j+d-2)+m$, we know that E_j is the space of restrictions to S^{d-1} of all harmonic polynomials on $\mathbb{R}^d$ homogeneous of degree j. Actually for $d=2$ (which corresponds to 1-d nonlinear wave equation with periodic boundary condition), E_j is the linear subspace spaned by e^{ijx} and e^{-ijx} and has the constant dimension two. For $d\geq 3$ the dimension of

E_j grows like j^{d-1} (see [**BGM71**] for a general reference on Laplace-Beltrami operators). Of course since the same frequency is now associated to different modes, condition (4.5) is no more satisfied. Nevertheless, if we denote by e_j the dimension of E_j and by $\phi_{j,l}$, $l=1,\dots,e_j$ an orthonormal basis of E_j then Delort and Szeftel have proved in [**DS04, DS05**] that there exists $\nu\geq 0$ such that for any $k\geq 1$ and any $N\geq 0$ there exists a constant $C>0$ such that for any $j\in\mathbb{N}^k$ and any l_n with $1\leq l_n\leq e_{j_n}$ $(n=1,\dots,k)$

$$\int_{S^{d-1}}\phi_{j_1,l_1}\dots\phi_{j_k,l_k}dx\leq C\frac{\mu(j)^{N+\nu}}{S(j)^N}.$$

This estimate is a generalisation of (5.7) which means that, generalizing the definition 4.2, the perturbation will belong to $\mathcal{T}^\nu$.

Generalized statement. — In this subsection we present a generalisation of theorem 4.7 motivated by the previous examples. We follow the presentation of section 4 and only focus on the new feature.

Fix for any $j\geq 1$ an integer $e_j\geq 1$. We consider the phase space $\mathcal{Q}_s=\mathcal{L}_s\times\mathcal{L}_s$ with

$$\mathcal{L}_s=\{(a_{j,l})_{j\geq 1,\,1\leq l\leq e_j}\mid\sum_{j\geq 1}|j|^{2s}\sum_{l=1}^{e_j}|a_{j,l}|^2<\infty\}$$

that we endow with the standart norm and the standart symplectic structure as for $\mathcal{P}_s$ in section 4.1. We then define for $(q,p)\in\mathcal{Q}_s$,

$$H_0(q,p)=\frac{1}{2}\sum_{j\geq 1}\sum_{l=1}^{e_j}\omega_{j,l}(q_{j,l}^2+p_{j,l}^2)$$

and for $j\geq 1$,

$$J_j(q,p)=\frac{1}{2}\sum_{l=1}^{e_j}q_{j,l}^2+p_{j,l}^2\ .$$

We assume that the frequencies $\omega_{j,l}$ are weakly non resonant in the following sense:

Definition 7.1. — *The vector of frequencies $(\omega_{j,l})_{j\geq 1,\,1\leq l\leq e_j}$ is weakly non resonant if for any $k\in\mathbb{N}$, there are $\gamma>0$ and $\alpha>0$ such that for any $j\in\bar{\mathbb{N}}^k$, for any l_n with $1\leq l_n\leq e_{j_n}$ $(n=1,\dots,k)$ and for any $1\leq i\leq k$, one has*

$$|\omega_{j_1,l_1}+\dots+\omega_{j_i,l_i}-\omega_{j_{i+1},l_{i+1}}-\dots-\omega_{j_k,l_k}|\geq\frac{\gamma}{\mu(j)^\alpha}\tag{7.2}$$

except if $\{j_1,\dots,j_i\}=\{j_{i+1},\dots,j_k\}$.

Notice that, with this definition, the frequencies of the same packet indexed by j (i.e. ω_{j_1,l_1} for $1 \leq l \leq e_j$) can be very close or even equal.
Using notations of section 4.1, we define the class $\mathbb{T}_k^{N,\nu}$ of real polynomials of degree k on $\mathcal{Q}_s$

$$Q(\xi,\eta) \equiv Q(z) = \sum_{m=0}^{k} \sum_{j\in\bar{\mathbb{Z}}^m} \sum_{l_1=1}^{e_{j_1}} \cdots \sum_{l_m=1}^{e_{j_m}} a_{j,l} z_{j_1,l_1} \ldots z_{j_m,l_m}$$

such that there exists a constant $C>0$ such that for all j,l

$$|a_{j,l}| \leq C \frac{\mu(j)^{N+\nu}}{S(j)^N}.$$

Then following definition 4.3 we define a corresponding class $\mathbb{T}^\nu$ of Hamiltonians on $\mathcal{Q}_s$ having a regular Hamiltonian vector field and Taylor's polynomials in $\mathbb{T}_k^{N,\nu}$.
Adapting the proof of theorem 4.7 we get

Theorem 7.2. — *Assume $P \in \mathbb{T}^\nu$ for some $\nu \geq 0$ and ω weakly non resonant in the sense of* (7.2). *Then for any $r \geq 3$ there exists s_0 and for any $s \geq s_0$ there exists $\mathcal{U}$, $\mathcal{V}$ neighborhoods of the origin in $\mathcal{Q}_s$ and $\tau : \mathcal{V} \to \mathcal{U}$ a real analytic canonical transformation which puts $H = H_0 + P$ in normal form up to order r i.e.*

$$H \circ \tau = H_0 + Z + R$$

with

(i) *Z is a continuous polynomial of degree r which commutes with all J_j, $j \geq 1$, i.e. $\{Z, J_j\} = 0$ for all $j \geq 1$.*
(ii) *$R \in C^\infty(\mathcal{V},\mathbb{R})$ and $\|X_R(q,p)\|_s \leq C_s \|(q,p)\|_s^r$ for all $(q,p) \in \mathcal{V}$.*
(iii) *τ is close to the identity: $\|\tau(q,p) - (q,p)\|_s \leq C_s \|(q,p)\|_s^2$ for all $(q,p) \in \mathcal{V}$.*

This theorem is an abstract version of theorem 2.6 in [**BDGS05**]. Notice that the concept of normal form is not the same as in theorem 4.7: the normal form $H_0 + Z$ is no more, in general, integrable. The dynamical consequences are the same as in corollary 4.8 but we have to replace I_j by J_j in the second assertion. Actually the J_j play the rule of almost actions: they are almost conserved quantities.
This abstract theorem applies to both examples that we present at the begining of this section and thus the dynamical corollary also. For a proof, rafinements and comments, see [**BG04**] for the case of periodic boudary conditions and [**BDGS05**] for the case of the Klein-Gordon equation on the sphere. Notice

that, in this last context, the fact that Z commutes with all the J_j can be interpreted saying that $H_0 + Z$ only allows energy exchanges between modes in the same packet E_j (i.e. that correspond to the same frequency).

We finally notice that in [**BDGS05**], the normal form was used to prove an almost global existence result for Klein-Gordon equations with small Cauchy data on the sphere (and more generally on Zoll manifold).

7.2. Comments on KAM theory. — In this section we briefly introduce the KAM theory in finite dimension and then we give an idea of the (partial) generalisation to the infinite dimensional case. Our aim is to compare these results to the Birkhoff approach developped in these notes.
For a simple introduction to the KAM theory in finite dimension we refer to [**Way96**] and [**HI04**] (which both include a complete proof of KAM theorem) and to the second chapter of [**KP03**]. For infinite dimensional context, the reader may consult the books by S. Kuksin [**Kuk93, Kuk00**] or the one by T. Kappeler and J. Pöschel [**KP03**].

The classical KAM theorem. — In contrast with section 3 we consider Hamiltonian perturbations of *Liouville* integrable system: $H = H_0 + P$. We denote by[7] $(I, \theta) \in \mathbb{R}^n \times T^n$ the action-angle variables for H_0 and ω_j, $j = 1, \dots, n$, the free frequencies. One has $\omega_j = \frac{\partial H}{\partial I_j}$ and the unperturbed equations read

$$\begin{cases} \dot{I}_j = 0, & j = 1, \dots, n \ , \\ \dot{\theta}_j = \omega_j, & j = 1, \dots, n \ . \end{cases}$$

The phase space $M = \mathcal{U} \times T^n$, where $\mathcal{U}$ is an open bounded domain of $\mathbb{R}^n$, is foliated by the invariant tori

$$T_I = T^n \times \{I\}.$$

Our problem is to decide if these tori will persist after small hamiltonian perturbation of the system.
For simplicity, we assume that the perturbation is of the form $P = \epsilon F$. The Hamiltonian equation associated to H then read

$$\begin{cases} \dot{I}_j = -\epsilon \frac{\partial F}{\partial \theta_j}, & j = 1, \dots, n \ , \\ \dot{\theta}_j = \omega_j + \epsilon \frac{\partial F}{\partial I_j}, & j = 1, \dots, n \ . \end{cases} \tag{7.3}$$

To guarantee the persistency of T_I, it is not sufficient to assume the nondegenerancy of the frequencies (see definition 2.1) and we need the following

[7] here $T^n = S^1 \times \ldots \times S^1$, n times, is the n dimensional torus

Definition 7.3. — *A frequencies vector $\omega \in \mathbb{R}^n$ is diophantine if there exist constants $\gamma > 0$ and $\alpha > 0$ such that for all $0 \neq k \in \mathbb{Z}^n$*

$$|k \cdot \omega| \geq \frac{\gamma}{|k|^\alpha} \ . \tag{7.4}$$

We denote by D_γ the set of frequencies satisfying (7.4) for some $\alpha > 0$. It turns out that almost every vector in $\mathbb{R}^n$ are diophantine: by straighforward estimates one proves that the set of vectors in a bouded domain of $\mathbb{R}^n$ that do not belong to D_γ has Lebesgue measure $O(\gamma)$.

The second condition that we will need says that the frequencies effectively vary with the actions and thus we cannot stay in a resonnant situation when varying the actions:
The unperturbed system is said *nondegenerate* on $\mathcal{U}$ if the Hessian matrix of H_0

$$\mathrm{Hess}_{H_0}(I) = \left(\frac{\partial^2 H_0}{\partial I_j \partial I_k}(I)\right)_{1\leq j,k\leq n}$$

is invertible on $\mathcal{U}$. This nondegenerancy condition insures that the frequency map

$$I \mapsto \omega(I) = \left(\frac{\partial H_0}{\partial I_j}_{1\leq j\leq n}\right)$$

is a local diffeomorphism at each point of $\mathcal{U}$.
Notice that this condition is not satisfied by the harmonic oscillator, $H_0 = \sum \omega_j I_j$, for which the frequency map is constant. This makes difficult to directly compare theorem 3.2 and theorem 7.4 below.

Theorem 7.4. — *(The classical KAM theorem* [**Kol54, Arn63, Mos62**]*) Assume that $(I,\theta) \mapsto H = H_0 + \epsilon F$ is real analytic on the closure of $\mathcal{U} \times T^n$ and that H_0 is nondegenerate on $\mathcal{U}$. There exists $C > 0$ such that if $\epsilon \leq C\gamma^2$ and if $I \in \mathcal{U}$ is such that the frequencies vector $\omega(I)$ belongs to D_γ then the corresponding torus T_I persists after perturbation.*

As a dynamical consequence, we deduce that the system of equations (7.3) has a lot of quasiperiodic solutions. But to decide if an invariant torus T_I survives the perturbation, we have to know if the corresponding frequencies are in a Cantor type set. As we said in the introduction, this is not a realistic physical condition. That's why, even in the finite dimensional case, we can prefer to use the Birkhoff theory which provides long time stability under the condition that the frequencies are in an open subset of full Lebesgue measure.

The KAM theorem only concerns the Lagrangian tori, i.e. tori of maximal dimension. We can also wonder what happens to the lower dimensional tori. For instance if we fix the $n-m$ last actions to the value 0 then we can define angle variables only for the m first actions and the corresponding invariant torus T_I is diffeomorphic to $T^m \times \{0_{\mathbb{R}^{n-m}}\} \times \{I\}$ whose dimension is $m < n$. This difficult problem has been solved by H. Elliasson [**Eli88**] under the so called Melnikov condition [(8)] which says that, as a function of the first m actions denoted by $\tilde{I}$, the quantities

$$\sum_{j=1}^{m} k_j \omega_j(\tilde{I}) + \sum_{j=m+1}^{n} l_j \omega_j(\tilde{I})$$

does not vanish identically (and thus effectively vary with $\tilde{I}$ since H is real analytic) for all non trivial $(k,l) \in \mathbb{Z}^m \times \mathbb{Z}^{n-m}$ with $|l| \le 2$.
The theorem then says, roughly speaking, that under the hypothesis that H_0 is non degenerate and satisfies the Melnikov condition, for sufficiently small values of ϵ, there exists a Cantor set of effective actions $\tilde{I}$ for which the corresponding invariant tori survive the small perturbation (cf. [**Eli88**] or [**KP03**] for a precise statement).

Extension to the infinite dimensional case. — When trying to extend theorem 7.4 to the infinite dimensional case, we face, as in the case of Birkhoff theorem, the problem of extending the nonresonancy condition. It turns out that, because of the Dirichlet's theorem, the condition (7.4) cannot be satisfied for all k when the number of frequencies involved grows to infinity. So we cannot expect a polynomial control of the small divisors and it is very difficult to preserve tori of infinite dimension. In PDEs context, this would imply the existence of almost periodic solution, i.e. quasi-periodic solutions with a frequencies vector of infinite dimension. Unfortunately, up to now, there is essentially no result in this direction (see however the recent result by J. Bourgain [**Bou05b**]). The only case where there exists a result applying to realistic PDEs concerns the perturbation of *finite dimensional tori.* Of course, the set of finite dimensional tori is very small within an infinite dimensional phase space, but it allows to describe the quasiperiodic solutions which is already very interesting.
A finite dimensional torus in an infinite dimensional phase space plays the role of a lower dimensional torus in a finite dimensional phase space and thus, it is not surprising that the crucial hypothesis in order to preserve a torus T_I of

[(8)] Actually V. K.Melnikov announced the result in [**Mel65**].

dimension N is a Melnikov condition:

$$\left|\sum_{j=1}^{N} k_j\omega_j + \sum_{j=N+1}^{\infty} l_j\omega_j\right| \geq \frac{\gamma}{|k|^{\alpha}} \tag{7.5}$$

for all $(k,l) \in \mathbb{Z}^N \times \mathbb{Z}^\infty$ with $|l| \leq 2$. The big differnce is that, now, the number of external frequencies, ω_j for $j \geq N+1$, is infinite. We are not trying to state a precise result in this direction, but it turns out that this Melnikov condition can be verified in certain PDE context (cf. [**Kuk93, Kuk00**] for precise statements and further references).

We would like to conclude these lectures with a comparison of this nonresonant condition with the condition introduced in definition 4.4. We remark that (4.5) can be written in the equivalent form

$$\left|\sum_{j=1}^{N} k_j\omega_j + \sum_{j=N+1}^{\infty} l_j\omega_j\right| \geq \frac{\gamma}{N^{\alpha}} \tag{7.6}$$

for all nontrivial $(k,l) \in \mathbb{Z}^N \times \mathbb{Z}^\infty$ with $|k| \leq r$ and $|l| \leq 2$.

Thus, (7.5) and (7.6) give a control of essentially the same type of small divisor but, in (7.5), N (the dimension of the torus that we perturb) is fixed and $|k|$ (the lenght of the divisor that we consider) is free while, in (4.7), $|k|$ (the degree of the monomials that we want to kill) is less than a fix r and N (the number of excited modes) is free.

References

[AG91] Serge Alinhac and Patrick Gérard, *Opérateurs pseudo-différentiels et théorème de Nash-Moser*, Savoirs Actuels. [Current Scholarship], InterEditions, Paris, 1991.

[Arn63] V. I. Arnold, *Proof of a theorem of A.N.Kolmogorov on the conservation of quasiperiodic motions under a small change of the Hamiltonian function*, Russ. Math. Surv. **18** (1963), no. 5, 9–36.

[Arn89] V. I. Arnold, *Mathematical methods in classical mechanics*, 3rd ed., Springer-Verlag, Berlin, 1989.

[Bam03] D. Bambusi, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Physics **234** (2003), 253–283.

[BDGS05] D. Bambusi, J. M. Delort, B. Grébert, and J. Szeftel, *Almost global existence for Hamiltonian semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds*, preprint (2005).

[BG04] D. Bambusi and B. Grébert, *Birkhoff normal form for PDEs with tame modulus*, To appear in Duke Math. J. (2004).

[BGG85] G. Benettin, L. Galgani, and A. Giorgilli, *A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems*, Celestial Mech. **37** (1985), 1–25.

[BGM71] Marcel Berger, Paul Gauduchon, and Edmond Mazet, *Le spectre d'une variété riemannienne*, Lecture Notes in Mathematics, Vol. 194, Springer-Verlag, Berlin, 1971.

[Bou96] J. Bourgain, *Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations*, Geometric and Functional Analysis **6** (1996), 201–230.

[Bou05a] J. Bourgain, *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies, vol. 158, Princeton University Press, Princeton, NJ, 2005.

[Bou05b] J. Bourgain, *On invariant tori of full dimension for 1D periodic NLS*, J. Funct. Anal. **229** (2005), no. 1, 62–94.

[Cra00] W. Craig, *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthéses, no. 9, Société Mathématique de France, 2000.

[CW93] W. Craig and C. E. Wayne, *Newton's method and periodic solutions of nonlinear wave equations*, Comm. Pure Appl. Math. **46** (1993), 1409–1498.

[DS04] J. M. Delort and J. Szeftel, *Long–time existence for small data nonlinear Klein–Gordon equations on tori and spheres*, Internat. Math. Res. Notices **37** (2004), 1897–1966.

[DS05] J. M. Delort and J. Szeftel, *Long–time existence for semi–linear Klein–Gordon equations with small cauchy data on Zoll manifolds*, to appear in Amer. J. Math (2005).

[Eli88] L. H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **15** (1988), no. 1, 115–147 (1989).

[GG85] L. Galgani and A. Giorgilli, *Rigourous estimates for the series expansions of Hamiltonian perturbation theory*, Celestial Mech. **37** (1985), 95–112.

[HI04] John Hubbard and Yulij Ilyashenko, *A proof of Kolmogorov's theorem*, Discrete Contin. Dyn. Syst. **10** (2004), no. 1-2, 367–385, Partial differential equations and applications.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser, Basel, 1994.

[Kol54] A. N. Kolmogorov, *On the conservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk SSSR **98** (1954), 527–530.

[KP03] T. Kappeler and J. Pöschel, *KAM & KdV*, Springer, 2003.

[Kuk93] S. B. Kuksin, *Nearly integrable infinite-dimensional Hamiltonian systems*, Springer-Verlag, Berlin, 1993.

[Kuk00] S. B. Kuksin, *Analysis of Hamiltonian PDEs*, Oxford University Press, Oxford, 2000.

[Mar86] V. Marchenko, *Sturm-Liouville operators and applications*, Birkhäuser, Basel, 1986.

[Mel65] V. K. Melnikov, *On some cases of conservation of conditionally periodic motions under a small change of the hamilton function*, Soviet Math. Doklady **6** (1965), 1592–1596.

[Mos62] J. Moser, *On invariant curves of area preserving mappings of an annulus*, Nach. Akad. Wiss. Gött., Math. Phys. Kl. (1962), 1–20.

[Mos68] J. Moser, *Lectures on hamiltonian systems*, Mem. Amer. Math. Soc. **81** (1968), 1–60.

[MS71] J. Moser and C. L. Siegel, *Lectures on celestial mechanics*, Springer, Berlin, 1971.

[Nek77] N. N. Nekhoroshev, *Exponential estimate of the stability of near integrable Hamiltonian systems*, Russ. Math. Surveys **32** (1977), no. 6, 1–65.

[Pös93] J. Pöschel, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, Math. Z. **213** (1993), 187–216.

[PT87] J. Pöschel and E. Trubowitz, *Inverse spectral theory*, Academic Press, Boston, 1987.

[Way96] C. Eugene Wayne, *An introduction to KAM theory*, Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994), Lectures in Appl. Math., vol. 31, Amer. Math. Soc., Providence, RI, 1996, pp. 3–29.

Benoît Grébert

Laboratoire de Mathématique Jean Leray UMR 6629,

Université de Nantes,

2, rue de la Houssinière,

44322 Nantes Cedex 3, France

E-mail: benoit.grebert@univ-nantes.fr
